

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:02:27 ; Search time 42 Seconds
(without alignments)
600,974 Million cell updates/sec

Title: US-09-634-109d-2
Perfect score: 1634
Sequence: 1 MGRWVNSQSYDGFLLGIFS.....GEVGMALRKGLDRICRIGSOH 315

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------------------------|
| 1 | 1432 | 87.6 | 315 | 10 | US-09-886-055-175 Sequence 175, App |
| 2 | 880.5 | 53.9 | 369 | 10 | US-09-886-055-457 Sequence 457, App |
| 3 | 879.5 | 53.8 | 316 | 10 | US-09-886-055-477 Sequence 477, App |
| 4 | 869 | 53.2 | 316 | 10 | US-09-826-508-24 Sequence 24, App |
| 5 | 858 | 52.5 | 324 | 10 | US-09-886-055-429 Sequence 429, App |
| 6 | 846 | 51.8 | 317 | 10 | US-09-886-053-479 Sequence 479, App |
| 7 | 815 | 49.9 | 314 | 10 | US-09-761-288-16 Sequence 16, App |
| 8 | 815 | 49.9 | 314 | 10 | US-09-886-055-143 Sequence 143, App |
| 9 | 815 | 49.9 | 317 | 10 | US-09-761-288-18 Sequence 18, App |
| 10 | 814 | 49.8 | 312 | 10 | US-09-886-055-421 Sequence 421, App |
| 11 | 813.5 | 49.8 | 348 | 10 | US-09-886-055-433 Sequence 433, App |
| 12 | 810 | 49.6 | 305 | 10 | US-09-761-288-73 Sequence 73, App |
| 13 | 810 | 49.6 | 305 | 10 | US-09-761-288-75 Sequence 75, App |
| 14 | 810 | 49.6 | 305 | 10 | US-09-761-288-79 Sequence 79, App |
| 15 | 810 | 49.6 | 315 | 10 | US-09-886-055-435 Sequence 435, App |
| 16 | 798 | 48.8 | 338 | 10 | US-09-886-055-427 Sequence 427, App |
| 17 | 781 | 47.8 | 312 | 10 | US-09-886-055-473 Sequence 473, App |
| 18 | 762 | 46.6 | 331 | 10 | US-09-886-055-475 Sequence 475, App |
| 19 | 746.5 | 45.7 | 323 | 10 | US-09-886-055-163 Sequence 163, App |

| | | | | | |
|----|-------|------|-----|----|---------------------------------------|
| 20 | 741.5 | 45.4 | 348 | 10 | US-09-886-055-165 Sequence 165, App |
| 21 | 708 | 43.3 | 310 | 10 | US-09-747-835A-63 Sequence 63, App |
| 22 | 702 | 43.0 | 353 | 10 | US-09-747-835A-47 Sequence 47, App |
| 23 | 700 | 42.8 | 311 | 10 | US-09-864-761-46639 Sequence 46639, A |
| 24 | 692 | 42.4 | 312 | 9 | US-09-974-591-16 Sequence 16, App |
| 25 | 692 | 42.4 | 312 | 9 | US-09-974-591-18 Sequence 18, App |
| 26 | 691 | 42.3 | 310 | 9 | US-09-800-321A-36 Sequence 36, App |
| 27 | 691 | 42.3 | 316 | 10 | US-09-886-055-179 Sequence 179, App |
| 28 | 691 | 42.3 | 357 | 9 | US-09-800-321A-37 Sequence 37, App |
| 29 | 690 | 42.2 | 314 | 10 | US-09-771-730-32 Sequence 32, App |
| 30 | 690 | 42.2 | 357 | 9 | US-09-800-321A-35 Sequence 35, App |
| 31 | 689 | 42.2 | 307 | 10 | US-09-771-730-112 Sequence 112, App |
| 32 | 689 | 42.2 | 314 | 10 | US-09-771-730-24 Sequence 24, App |
| 33 | 689 | 42.2 | 319 | 10 | US-09-886-055-387 Sequence 387, App |
| 34 | 689 | 42.2 | 330 | 10 | US-09-886-055-199 Sequence 199, App |
| 35 | 688 | 42.1 | 307 | 10 | US-09-771-730-106 Sequence 106, App |
| 36 | 688 | 42.1 | 314 | 10 | US-09-771-730-22 Sequence 22, App |
| 37 | 688 | 42.1 | 314 | 10 | US-09-771-730-36 Sequence 36, App |
| 38 | 687 | 42.0 | 305 | 10 | US-09-771-730-114 Sequence 114, App |
| 39 | 687 | 42.0 | 305 | 10 | US-09-771-730-144 Sequence 144, App |
| 40 | 687 | 42.0 | 314 | 10 | US-09-771-730-26 Sequence 26, App |
| 41 | 687 | 42.0 | 314 | 10 | US-09-771-730-34 Sequence 34, App |
| 42 | 686 | 42.0 | 304 | 10 | US-09-771-730-136 Sequence 136, App |
| 43 | 685 | 41.9 | 307 | 10 | US-09-771-730-107 Sequence 107, App |
| 44 | 685 | 41.9 | 307 | 10 | US-09-771-730-113 Sequence 113, App |
| 45 | 685 | 41.9 | 317 | 10 | US-09-771-730-108 Sequence 108, App |

ALIGNMENTS

RESULT 1
US-09-886-055-175
Sequence 175, App
Patent No. US20020132273A1
GENERAL INFORMATION:
APPLICANT: STREYER, LOBERT
APPLICANT: ZOZULYA, SERGEY
TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
FILE REFERENCE: 078003-0277150
CURRENT APPLICATION NUMBER: US/09/886,055
CURRENT FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: 60/213,812
NUMBER OF SEQ ID NOS: 522
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 175
LENGTH: 315
TYPE: PRT
ORGANISM: Homo sapiens
US-09-886-055-175

Query Match 87.68; Score 1432; DB 10; Length 315;
Best Local Similarity 87.08; Pred. 5.3e-126;
Matches 274; Conservative 13; Mismatches 28; Indels 0; Gaps 0;
QY 1 MGRWVNSQSYDGFLLGIFSHSOTDVLFSVVMVFPALCGNVLFLIYLDAGLTPM 60
1 METWVNSQSYDGFLLGIFSHSOTDVLFSVVMVFPALCGNVLFLIYMDLHLP 60
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
DB 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
61 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
QY 1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSGEGLLGMA 120
1 YFELSLSLMDLMVCNIVPMANFLSGRKSISFVCGIOLIGFVSLVSG

OY 241 TCSSHLTAATLEFGAAMFYLRPRRYRAPSBDKVASIFVTYLTPLMLNPLIYSLRNGEVMG 300
 DB 241 TCSSHLTAATLEFGAAMFYLRPRRYRAPSBDKVASIFVTYLTPLMLNPLIYSLRNGEVMG 300
 OY 301 ALRKGIDRCRIGSOH 315
 DB 301 ALRKGIDRCRIGSOH 315

RESULT 2

US-09-886-055-457

Sequence 457, Application US/09886055

Patent No. US20020132273A1
 GENERAL INFORMATION:
 APPLICANT: STRYER, LUBERT
 APPLICANT: ZOZULYA, SERGEY
 TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
 TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
 FILE REFERENCE: 078003-0277150
 CURRENT APPLICATION NUMBER: US/09/886, 055
 PRIOR FILING DATE: 2001-06-22
 PRIOR APPLICATION NUMBER: 60/213, 812
 NUMBER OF SEQ ID NOS: 522
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 457
 LENGTH: 369
 TYPE: PR1
 ORGANISM: Homo sapiens
 US-09-886-055-457

Query Match 53.9%; Score 880.5; DB 10; Length 369;

Best Local Similarity 53.8%; Pred. No. 1.8e-74; Matches 163; Conservative 56; Mismatches 83; Indels 1; Gaps 1;

OY 6 NOSYTDGFFLLGIFSHQTDVLFSAVMVFTVALCGNVLLIFLYLDAGLHPMPYFELS 65
 DB 56 NSTSTD-FTMGLEFNKRETSGLIFATISITFFALMANGVMIFLQDLDLHPMPYFELS 114
 OY 66 QLSLMDLMLVCNIVPKMAANFLSGRKSISFVCGGIQIGFVSLVSGEGLLGLMAYDHYV 125
 DB 115 HLSIDMAYISTYVPMKLVNVLDDQRTISFVCGTAQHFLTLTIVGAEFFLLGLMAYDHYV 174
 OY 126 AVSHPLHYPLMNORVCLQITGSSMAFGIIDGVYQVMAAMGLPYGSRSDHFECEVAL 185
 DB 175 AICNPLRYPLMRSRKMIIAGSFGSLDGLFPLTPTMSFPCNSREINHFCEAPAV 234
 OY 186 LKACADTSLEDTLFLACVFMILLPFSIIMASYACILGAVLRISAQAMKALATCSSH 245
 DB 235 LKACADTSLEDTLFLACVFMILLPFSIIMASYACILGAVLRISAQAMKALATCSSH 294
 OY 246 LTAATLEFGAAMFYLRPRRYRAPSBDKVASIFVTYLTPLMLNPLIYSLRNGEVMGALRKG 305
 DB 295 MTAVSLFYGAAMFYLRPRRYRAPSBDKVASIFVTYLTPLMLNPLIYSLRNGEVMGALRKA 354
 OY 306 LDR 308
 DB 355 LGR 357

RESULT 3

US-09-886-055-477

Sequence 477, Application US/09886055

Patent No. US20020132273A1
 GENERAL INFORMATION:
 APPLICANT: STRYER, LUBERT
 APPLICANT: ZOZULYA, SERGEY
 TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
 TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
 FILE REFERENCE: 078003-0277150
 CURRENT APPLICATION NUMBER: US/09/886, 055
 PRIOR FILING DATE: 2001-06-22
 PRIOR APPLICATION NUMBER: 60/213, 812

PRIOR FILING DATE: 2000-06-22
 NUMBER OF SEQ ID NOS: 522
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 477
 LENGTH: 316
 TYPE: PR1
 ORGANISM: Homo sapiens
 US-09-886-055-477

Query Match 53.8%; Score 879.5; DB 10; Length 316;

Best Local Similarity 53.6%; Pred. No. 1.9e-74; Matches 163; Conservative 59; Mismatches 81; Indels 1; Gaps 1;

OY 6 NOSYTDGFFLLGIFSHQTDVLFSAVMVFTVALCGNVLLIFLYLDAGLHPMPYFELS 65
 DB 3 NSTSTD-FTMGLEFNKRETSGLIFATISITFFALMANGVMIFLQDLDLHPMPYFELS 61
 OY 66 QLSLMDLMLVCNIVPKMAANFLSGRKSISFVCGGIQIGFVSLVSGEGLLGLMAYDHYV 125
 DB 62 QLSIMDTLICTVTPKLMADMSKEKISFVACGIQIFLYLTVMIGSEFFLLGLMAYDHYV 121
 OY 126 AVSHPLHYPLMNORVCLQITGSSMAFGIIDGVYQVMAAMGLPYGSRSDHFECEVAL 185
 DB 122 AYCNPRLRYPLMNRKCLLTAAGAMEGSLDGLFPLTPTMSFPCNSREINHFCEAPAV 181
 OY 186 LKACADTSLEDTLFLACVFMILLPFSIIMASYACILGAVLRISAQAMKALATCSSH 245
 DB 182 LKACADTSLEDTLFLACVFMILLPFSIIMASYACILGAVLRISAQAMKALATCSSH 241
 OY 246 LTAATLEFGAAMFYLRPRRYRAPSBDKVASIFVTYLTPLMLNPLIYSLRNGEVMGALRKG 305
 DB 242 LTVASIFYGAAFYTYVLPQSFHTPEODKVASAFYTTVTPMLNPLIYSLRNGEVMGALRKA 301
 OY 306 LDRG 309
 DB 302 FACC 305

RESULT 4

US-09-826-508-24

Sequence 24, Application US/09826508

Patent No. US20010025099A1
 GENERAL INFORMATION:
 APPLICANT: Nabil Elshourbagy
 TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
 FILE REFERENCE: GP-70744USB
 CURRENT APPLICATION NUMBER: US/09/826, 508
 CURRENT FILING DATE: 2001-04-05
 NUMBER OF SEQ ID NOS: 40
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 24
 LENGTH: 316
 TYPE: PR1
 ORGANISM: HOMO SAPIENS
 US-09-826-508-24

Query Match 53.2%; Score 869; DB 10; Length 316;

Best Local Similarity 54.2%; Pred. No. 1.8e-73; Matches 162; Conservative 47; Mismatches 90; Indels 0; Gaps 0;

OY 6 NOSYTDGFFLLGIFSHQTDVLFSAVMVFTVALCGNVLLIFLYLDAGLHPMPYFELS 65
 DB 5 NSTSTD-FTMGLEFNKRETSGLIFATISITFFALMANGVMIFLQDLDLHPMPYFELS 64
 OY 66 QLSLMDLMLVCNIVPKMAANFLSGRKSISFVCGGIQIGFVSLVSGEGLLGLMAYDHYV 125
 DB 65 QLSLMDLMLICTVTPKLMADMSKEKISFVACGIQIFLYLTVMIGSEFFLLGLMAYDHYV 124
 OY 126 AVSHPLHYPLMNORVCLQITGSSMAFGIIDGVYQVMAAMGLPYGSRSDHFECEVAL 185
 DB 125 AICHLRYTNLMSPRKICGLMTAFSWILGSTDGIYAATVTFSSYCGSRSEIAHFCELPSTL 184

US-09-761-288-16

Query Match 49.9%; Score 815; DB 10; Length 314;
 Best Local Similarity 54.1%; Pred. No. 2e-68;
 Matches 165; Conservative 48; Mismatches 90; Indels 2; Gaps 2;

QY 5 VNOSYTDGFFLLGIFSHSQDVLFSAVMVFYALCGNVLLIFLYLDAGLHPTMYEFL 64
 Db 4 VNOSVASFLLVGLFHSRSGROLFLSLVAWVFYGLGNVLLFLIRVDSRLHPTMYEFL 63
 QY 65 SOLSLMDLMLVNCNIPKMANFLSGRKSISFVCGCIQIGFVSIVG-SEGLLGLMAYDH 123
 Db 64 SOLSLFDIGCPMWTIPKMASDFLNGEGATSGGGAOI-FFLTLMGVAEGVLLVLMYDR 122
 QY 124 YVAVSHPLHPIILMNQVCIQITGSSWAFGIIDGVIOMVAMGLPYCGSSVDHFECEVQ 183
 Db 123 YVAVCOPLQYVLMRRQVCLLMGSSWVGVNLNSIOTSTILHPPYCASRIYDHFECEVP 182
 QY 184 ALKLACADTSLEFDTLLFACCVFMLLPFSITIMASYACIIGAVLRIRSAQMKALATCS 243
 Db 183 ALKLSCADTCAYEMALSTSGVLLTMLPLSLIATSYGHVLAVALMSRSEARHKAATYTC 242
 QY 244 SHLTAVTLFYGAAMFYLRRRYPAPSHDKVASIFYVTLPMNLPLYSLRNGEVMGALR 303
 Db 243 SHITVGLFYGAAMFYMVPAYHSPQDDNVVSLFYSLVPTLNPPLYSLRNPPEVMALV 302
 QY 304 KGLDR 308
 Db 303 KVLRS 307

RESULT 8

US-09-886-055-143
 ; Sequence 143, Application US/09886055
 ; Patent No. US20020132273A1
 ; GENERAL INFORMATION:
 ; APPLICANT: STREYER, LOBERT
 ; APPLICANT: 202017A, SERGEY
 ; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
 ; FILE REFERENCE: 078003-0277150
 ; CURRENT APPLICATION NUMBER: US/09/886,055
 ; CURRENT FILING DATE: 2001-06-22
 ; PRIOR APPLICATION NUMBER: 60/213,812
 ; PRIOR FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 522
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 143
 ; LENGTH: 314
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-886-055-143

Query Match 49.9%; Score 815; DB 10; Length 314;
 Best Local Similarity 54.1%; Pred. No. 2e-68;
 Matches 165; Conservative 48; Mismatches 90; Indels 2; Gaps 2;

QY 5 VNOSYTDGFFLLGIFSHSQDVLFSAVMVFYALCGNVLLIFLYLDAGLHPTMYEFL 64
 Db 4 VNOSVASFLLVGLFHSRSGROLFLSLVAWVFYGLGNVLLFLIRVDSRLHPTMYEFL 63
 QY 65 SOLSLMDLMLVNCNIPKMANFLSGRKSISFVCGCIQIGFVSIVG-SEGLLGLMAYDH 123
 Db 64 SOLSLFDIGCPMWTIPKMASDFLNGEGATSGGGAOI-FFLTLMGVAEGVLLVLMYDR 122
 QY 124 YVAVSHPLHPIILMNQVCIQITGSSWAFGIIDGVIOMVAMGLPYCGSSVDHFECEVQ 183
 Db 123 YVAVCOPLQYVLMRRQVCLLMGSSWVGVNLNSIOTSTILHPPYCASRIYDHFECEVP 182
 QY 184 ALKLACADTSLEFDTLLFACCVFMLLPFSITIMASYACIIGAVLRIRSAQMKALATCS 243
 Db 183 ALKLSCADTCAYEMALSTSGVLLTMLPLSLIATSYGHVLAVALMSRSEARHKAATYTC 242

QY 244 SHLTAVTLFYGAAMFYLRRRYPAPSHDKVASIFYVTLPMNLPLYSLRNGEVMGALR 303
 Db 243 SHITVGLFYGAAMFYMVPAYHSPQDDNVVSLFYSLVPTLNPPLYSLRNPPEVMALV 302
 QY 304 KGLDR 308
 Db 303 KVLRS 307

RESULT 9

US-09-761-288-18
 ; Sequence 18, Application US/09761288
 ; Patent No. US20020065405A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Prayaga, Sudhidas
 ; APPLICANT: Taupler, Raymond J
 ; APPLICANT: Mishra, Vishnu
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Li, Li
 ; TITLE OF INVENTION: NO. US20020065405A1el Polypeptides and Nucleic Acids Encoding
 ; FILE REFERENCE: 15966-638
 ; CURRENT APPLICATION NUMBER: US/09/761,288
 ; CURRENT FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: 60/177,839
 ; PRIOR FILING DATE: 2000-01-25
 ; PRIOR APPLICATION NUMBER: 60/176,134
 ; PRIOR FILING DATE: 2000-01-14
 ; PRIOR APPLICATION NUMBER: 60/175,989
 ; PRIOR FILING DATE: 2000-01-13
 ; PRIOR APPLICATION NUMBER: 60/218,324
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: 60/220,253
 ; PRIOR FILING DATE: 2000-07-24
 ; PRIOR APPLICATION NUMBER: 60/178,191
 ; PRIOR FILING DATE: 2000-01-26
 ; PRIOR APPLICATION NUMBER: 60/178,227
 ; PRIOR FILING DATE: 2000-01-26
 ; PRIOR APPLICATION NUMBER: 60/220,590
 ; PRIOR FILING DATE: 2000-07-25
 ; NUMBER OF SEQ ID NOS: 95
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 18
 ; LENGTH: 317
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-761-288-18

Query Match 49.9%; Score 815; DB 10; Length 317;
 Best Local Similarity 54.1%; Pred. No. 2e-68;
 Matches 165; Conservative 48; Mismatches 90; Indels 2; Gaps 2;

QY 5 VNOSYTDGFFLLGIFSHSQDVLFSAVMVFYALCGNVLLIFLYLDAGLHPTMYEFL 64
 Db 4 VNOSVASFLLVGLFHSRSGROLFLSLVAWVFYGLGNVLLFLIRVDSRLHPTMYEFL 63
 QY 65 SOLSLMDLMLVNCNIPKMANFLSGRKSISFVCGCIQIGFVSIVG-SEGLLGLMAYDH 123
 Db 64 SOLSLFDIGCPMWTIPKMASDFLNGEGATSGGGAOI-FFLTLMGVAEGVLLVLMYDR 122
 QY 124 YVAVSHPLHPIILMNQVCIQITGSSWAFGIIDGVIOMVAMGLPYCGSSVDHFECEVQ 183
 Db 123 YVAVCOPLQYVLMRRQVCLLMGSSWVGVNLNSIOTSTILHPPYCASRIYDHFECEVP 182
 QY 184 ALKLACADTSLEFDTLLFACCVFMLLPFSITIMASYACIIGAVLRIRSAQMKALATCS 243
 Db 183 ALKLSCADTCAYEMALSTSGVLLTMLPLSLIATSYGHVLAVALMSRSEARHKAATYTC 242
 QY 244 SHLTAVTLFYGAAMFYLRRRYPAPSHDKVASIFYVTLPMNLPLYSLRNGEVMGALR 303
 Db 243 SHITVGLFYGAAMFYMVPAYHSPQDDNVVSLFYSLVPTLNPPLYSLRNPPEVMALV 302

| | | | |
|----|-----|-----|-----|
| QY | 304 | KGL | 306 |
| | | | |
| Db | 303 | KVL | 305 |

| | | | | |
|-----------------------|------------------|--------------------|-----------|-------------|
| Query Match | 49.6%; | Score 810; | DB 10; | Length 305; |
| Best Local Similarity | 54.1%; | Pred. No. 5.6e-68; | | |
| Matches 164; | Conservative 48; | Mismatches 89; | Indels 2; | Gaps 2 |

| | | | |
|----|-----|-----|-----|
| QY | 304 | KGL | 306 |
| | | | |
| Db | 303 | KVL | 305 |

| | | | | | | | |
|-----------------------|-------|--------------|-----|------------|----|--------|-----|
| Query Match | 49.6% | Score | 810 | DB | 10 | length | 305 |
| Best Local Similarity | 54.1% | Pred. No. | 5 | 6e-68 | | | |
| Matches | 164 | conservative | 48 | Mismatches | 89 | Indels | 2 |
| | | | | Gaps | 2 | | |

| | | | |
|----|---|---|----|
| QY | 5 | VNOSTDGEFLGIGTSHSOTDLVFSANVYVTVALCGVALLFLIYDAGHTPMYFYL | 64 |
| | | | |
| DB | 4 | VNVAASDILGVLGSHSGSRLQLFSLVAMKIVIGLIGATVLLFLRDSRHTPMYFYL | 63 |
| | | | |

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:52:36 ; Search time 39 Seconds

(without alignments)
1076.255 Million cell updates/sec

Title: US-09-634-109d-2

Perfect score: 1634
Sequence: 1 MGRMVQSYDTGFLLGIFS.....GEVWGALRKGLDRCRIGSOH 315

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-----------------------------|
| 1 | 1634 | 100.0 | 315 | 22 | AAU00529 Human CON167 G pro |
| 2 | 1626 | 99.5 | 326 | 22 | AAU00529 Human olfactory re |
| 3 | 1620 | 99.1 | 315 | 23 | ABG68141 Human G-protein co |
| 4 | 1432 | 87.6 | 315 | 22 | AAU01699 Human olfactory re |
| 5 | 1432 | 87.6 | 315 | 22 | AAU01699 Human olfactory re |
| 6 | 1432 | 87.6 | 315 | 22 | AAU01699 Human olfactory re |
| 7 | 1432 | 87.6 | 315 | 22 | AAU01699 Human olfactory re |
| 8 | 1432 | 87.6 | 315 | 23 | AAU01699 Human olfactory an |
| 9 | 1432 | 87.6 | 315 | 23 | AAU01699 Human G-protein co |
| 10 | 1360 | 83.2 | 271 | 23 | AAU01699 Human G-protein co |

| | | | | | |
|----|-------|------|-----|----|-----------|
| 11 | 1150 | 70.4 | 252 | 22 | AAU024603 |
| 12 | 1146 | 70.1 | 263 | 23 | ABP61132 |
| 13 | 1044 | 63.9 | 223 | 21 | AAU06666 |
| 14 | 1044 | 63.9 | 223 | 21 | AAU06667 |
| 15 | 1040 | 63.6 | 223 | 21 | AAU06668 |
| 16 | 911 | 55.8 | 320 | 23 | AAU095784 |
| 17 | 902 | 55.2 | 311 | 23 | AAU095780 |
| 18 | 896 | 54.8 | 320 | 23 | AAU095782 |
| 19 | 894 | 54.0 | 312 | 23 | AAU095781 |
| 20 | 882.5 | 54.0 | 316 | 22 | ABB81453 |
| 21 | 880.5 | 53.9 | 311 | 22 | AAU07173 |
| 22 | 880.5 | 53.9 | 318 | 23 | AAU095679 |
| 23 | 880.5 | 53.9 | 369 | 22 | AAU024744 |
| 24 | 880.5 | 53.9 | 369 | 23 | AAU085364 |
| 25 | 879.5 | 53.8 | 316 | 22 | AAU024754 |
| 26 | 879.5 | 53.8 | 316 | 23 | ABB81452 |
| 27 | 879.5 | 53.8 | 316 | 23 | AAU085374 |
| 28 | 878 | 53.7 | 312 | 23 | AAU095783 |
| 29 | 870 | 53.2 | 289 | 23 | AAU095669 |
| 30 | 869 | 53.2 | 316 | 19 | AAU075960 |
| 31 | 862 | 52.8 | 321 | 22 | AAU071467 |
| 32 | 858 | 52.5 | 324 | 22 | AAU024730 |
| 33 | 858 | 52.5 | 324 | 23 | AAU095680 |
| 34 | 858 | 52.5 | 324 | 23 | AAU085350 |
| 35 | 858 | 52.5 | 324 | 23 | AAU085350 |
| 36 | 849 | 52.0 | 317 | 22 | ABB81447 |
| 37 | 846 | 51.8 | 317 | 23 | AAU024755 |
| 38 | 846 | 51.8 | 317 | 23 | ABB81446 |
| 39 | 846 | 51.8 | 317 | 23 | AAU085375 |
| 40 | 843 | 51.6 | 318 | 22 | AAU085375 |
| 41 | 840 | 51.4 | 317 | 23 | ABB81455 |
| 42 | 840 | 51.4 | 317 | 23 | ABB81456 |
| 43 | 836 | 51.2 | 317 | 23 | ABP51586 |
| 44 | 828 | 50.7 | 312 | 23 | AAU085285 |
| 45 | 826 | 50.6 | 275 | 23 | ABB81454 |

ALIGNMENTS

| | | |
|----------|--|---------------------------------------|
| RESULT 1 | AAU00529 | standard; Protein: 315 AA. |
| ID | AAU00529 | |
| XX | AAU00529: | |
| XX | 14-MAY-2001 | (first entry) |
| XX | Human CON167 G protein seven transmembrane receptor. | |
| XX | CON167: G protein coupled receptor; transmembrane receptor. | |
| KW | neurological disorder; psychiatric disease; schizophrenia; depression; | |
| KW | anxiety; bipolar disorder; affective disorder; epilepsy; neuritis; | |
| KW | attention deficit hyperactivity disorder; ADHD; neuroasthenia; neuropathy; | |
| KW | neurosis; Alzheimer's disease; Parkinson's disease; migraine; therapy; | |
| KW | senile dementia; hybridisation assay; diagnosis; transgenic animal. | |
| XX | Homo sapiens. | |
| OS | Homo sapiens. | |
| XX | Key | Location/Qualifiers |
| XX | Key | 27..51 |
| XX | Domain | /note="Transmembrane-spanning domain" |
| XX | Domain | 60..79 |
| XX | Domain | /note="Transmembrane spanning-domain" |
| XX | Domain | 92..121 |
| XX | Domain | /note="Transmembrane spanning-domain" |
| XX | Domain | 151..170 |
| XX | Domain | /note="Transmembrane spanning-domain" |
| XX | Domain | 196..220 |
| XX | Domain | /note="Transmembrane spanning-domain" |
| XX | Domain | 242..260 |
| XX | Domain | /note="Transmembrane spanning-domain" |
| XX | Domain | 274..294 |

/note- "Transmembrane spanning-domain"

WT 00000114554-A1.

XX 01-MAR-2001.

XX 08-AUG-2000; 2000MO-US21566.

XX 19-AUG-1999; 99US-0377563.

XX (PHAA) PHARMACIA & UPJOHN CO.

XX Vogel I G, Wood LS:

XX WPI: 2001-218450/22.

XX N-PSDB; AAS00555.

XX Novel purified isolated seven transmembrane receptor polypeptide

XX (G-protein coupled receptor) useful for treating neurological and

XX psychiatric diseases such as schizophrenia, depression, anxiety,

XX bipolar disease and affective disorder

XX Claim 1; Page 63-64; 72pp; English.

XX The sequence represents human CON167 G protein seven transmembrane

XX receptor. The protein and its corresponding DNA are useful for

XX modulating activity of CON167 in a mammal comprising cells that express

XX CON167, preferably in a human suffering from a neurological disorder

XX and/or psychiatric diseases such as schizophrenia, depression, anxiety,

XX bipolar disease, affective disorder, attention deficit hyperactivity

XX disorder (ADHD), epilepsy, neuritis, neurosclerosis, neuropathy, neuroses,

XX Alzheimer's disease, Parkinson's disease, migraine and senile dementia.

XX The polynucleotides are useful in hybridisation assays to detect the

XX capacity of cells to express CON167, for large scale expression of CON167

XX required for research, for identification and isolation of new

XX polynucleotides encoding the related CON167 polypeptides, in diagnostic

XX methods for identifying any genetic alteration in a CON167 locus that

XX underlies a disease state (which is useful for selection of therapeutic

XX strategies), and for the development of transgenic animals that fail to

XX express functional CON167 or that express a variant of CON167.

XX

XX Sequence 315 AA:

XX

XX Query Match 100.0%; Score 1634; DB 22; Length 315;

XX Best Local Similarity 100.0%; Pred. No. 9,8e-179;

XX Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

XX 1 MGRVWNSYTDGFFLLGIFSHSQTDVLFSAVMVFTVALCGNLLFLFLYLDAGLHTPM 60

XX

XX 61 YFFLSQSLMDMLVNCINVPKMANFLSGRKSISFVCGIOIGFVSLVSGEGLLGIMA 120

XX

XX 61 YFFLSQSLMDMLVNCINVPKMANFLSGRKSISFVCGIOIGFVSLVSGEGLLGIMA 120

XX

XX 121 YDHYVAVSHPLHYPIILNNQRYCQITGSSNAFGIIDGVIOQVAAVAMGLPYCGSRSDVDFEC 180

XX

XX 121 YDHYVAVSHPLHYPIILNNQRYCQITGSSNAFGIIDGVIOQVAAVAMGLPYCGSRSDVDFEC 180

XX

XX 181 EVOALLKLCACADTSLFDLTFACCVFMLLPFSITMASYACILGAVLRIRSAQAMKRALA 240

XX

XX 181 EVOALLKLCACADTSLFDLTFACCVFMLLPFSITMASYACILGAVLRIRSAQAMKRALA 240

XX

XX 241 TCSSHLTAVTLEFYGAAMFMYLRPRRYRAPSHDKVASIFYYTLVPLMLNPLIYSLNAGEVMG 300

XX

XX 241 TCSSHLTAVTLEFYGAAMFMYLRPRRYRAPSHDKVASIFYYTLVPLMLNPLIYSLNAGEVMG 300

XX

XX 301 ALRKGDLRCRIGSOH 315

XX

XX 301 ALRKGDLRCRIGSOH 315

XX

XX RESULT 2

AAG71793

ID AAG71793 standard; Protein; 326 AA.

XX AAG71793;

XX 31-JUL-2001 (first entry)

XX

XX Human olfactory receptor polypeptide, SEQ ID NO: 1474.

XX

XX Human; olfactory receptor; OR; primary scent determination;

XX secondary scent determination; polypeptide library; odour receptor;

XX scent profile; scent fingerprint; scent representation.

XX

XX Homo sapiens.

XX

XX W0200127158-A2.

XX

XX 19-APR-2001.

XX

XX 06-OCT-2000; 2000MO-US27582.

XX

XX 08-OCT-1999; 99US-0158615.

XX

XX 24-FEB-2000; 2000US-0184809.

XX

XX (DIGI-) DIGISCENTS.

XX (YEDA) YEDA RES & DEV CO LTD.

XX

XX Ballenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

XX WPI: 2001-290713/30.

XX

XX New polynucleotides which encode polypeptides involved in olfactory

XX sensation for identifying olfactory agonists and antagonists -

XX

XX Claim 11; Page 932-933; 1857pp; English.

XX

XX The present sequence is an olfactory receptor which is encoded by

XX one of a number of novel polynucleotides. The polynucleotides can be

XX used in screening for olfactory agonists and antagonists. The methods

XX allow for the determination of primary scents and the identification

XX of the odour receptors used to detect these primary scents. The methods

XX also enable determination of secondary scents and the identification of

XX combinations of odour receptors that are involved in detecting such

XX secondary scents. This enables the construction of a scent representation

XX (also called a scent fingerprint or scent profile), which may be used to

XX re-create and edit scents. Libraries of olfactory receptors are useful

XX for determining the interaction pattern of a composition with the

XX receptors, and can be used for determining differences in the olfactory

XX faculties of different individuals.

XX

XX Sequence 326 AA:

XX

XX Query Match 99.5%; Score 1626; DB 22; Length 326;

XX Best Local Similarity 99.7%; Pred. No. 8,5e-178;

XX Matches 314; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX

XX 1 MGRVWNSYTDGFFLLGIFSHSQTDVLFSAVMVFTVALCGNLLFLFLYLDAGLHTPM 60

XX

XX 61 YFFLSQSLMDMLVNCINVPKMANFLSGRKSISFVCGIOIGFVSLVSGEGLLGIMA 120

XX

XX 61 YFFLSQSLMDMLVNCINVPKMANFLSGRKSISFVCGIOIGFVSLVSGEGLLGIMA 120

XX

XX 121 YDHYVAVSHPLHYPIILNNQRYCQITGSSNAFGIIDGVIOQVAAVAMGLPYCGSRSDVDFEC 180

XX

XX 121 YDHYVAVSHPLHYPIILNNQRYCQITGSSNAFGIIDGVIOQVAAVAMGLPYCGSRSDVDFEC 180

XX

XX 181 EVOALLKLCACADTSLFDLTFACCVFMLLPFSITMASYACILGAVLRIRSAQAMKRALA 240

XX

XX 181 EVOALLKLCACADTSLFDLTFACCVFMLLPFSITMASYACILGAVLRIRSAQAMKRALA 240

XX

XX 241 TCSSHLTAVTLEFYGAAMFMYLRPRRYRAPSHDKVASIFYYTLVPLMLNPLIYSLNAGEVMG 300

DB 241 TCSSHLTAVTLEFYGAAMFMYLRPRRYRAPSHDKVASIFTYVLPMLNPLIYSLRNGEVMG 300
 QY 301 ALRKGDLRCRIGSOH 315
 DB 301 ALRKGDLRCRIGSOH 315
 RESULT 3
 ABG68141
 ID ABG68141 standard; Protein; 315 AA.
 XX ABG68141;
 AC
 XX
 DT 07-OCT-2002 (first entry)
 DE Human G-protein coupled receptor (GPCR) GPCR protein #8.
 XX
 KW G-protein coupled receptor; receptor; GPCR; GPCR; cardiomypathy;
 KW atherosclerosis; diabetes; cell signal processing; cancer; trauma;
 KW metabolic pathway modulation; neuro-olfactory system; surgery;
 KW neoplastic disorder; adenocarcinoma; lymphoma; prostate cancer;
 KW uterine cancer; immune response; acquired immunodeficiency syndrome;
 KW AIDS; asthma; Crohn's disease; multiple sclerosis;
 KW Albright hereditary osteodystrophy.
 XX
 OS Homo sapiens.
 XX
 PN WO200250117-A2.
 XX
 PD 27-JUN-2002.
 XX
 PF 18-DEC-2001; 2001MO-US49077.
 XX
 PR 18-DEC-2000; 2000US-256635P.
 PR 21-DEC-2000; 2000US-257876P.
 PR 04-JAN-2001; 2001US-259743P.
 PR 10-JAN-2001; 2001US-260718P.
 PR 12-JAN-2001; 2001US-261498P.
 PR 24-JAN-2001; 2001US-263689P.
 PR 08-FEB-2001; 2001US-267464P.
 PR 22-FEB-2001; 2001US-271021P.
 PR 14-MAR-2001; 2001US-275946P.
 PR 23-MAR-2001; 2001US-278150P.
 PR 18-APR-2001; 2001US-284591P.
 PR 23-APR-2001; 2001US-285718P.
 PR 19-JUN-2001; 2001US-299327P.
 PR 16-AUG-2001; 2001US-312902P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Padigaru M, Kakuda R, Colman SD, Solytek KA, Ballinger RA;
 PI Vernet CAM, LI L, Shenoy S, Casman SJ;
 XX
 DR WPI: 2002-528447/56.
 DR N-PSDB; ABK97214.
 XX
 PT New G-protein coupled receptor polypeptides for treating or preventing
 PT cardiomypathy, atherosclerosis, diabetes, multiple sclerosis, acquired
 PT immunodeficiency syndrome or cancer in humans
 XX
 PS Claim 1: Page 86; 110pp; English.
 XX
 CC The present invention relates to a new G-protein coupled receptor (GPCR)
 CC polypeptide. The GPCR polypeptide, GPCR nucleic acid and antibody are
 CC useful for treating, preventing or alleviating a GPCR-associated
 CC disorder or a pathological state in a subject, particularly a human. In
 CC particular, the disorder is cardiomypathy, atherosclerosis, diabetes, or
 CC a disorder related to cell signal processing and metabolic pathway
 CC modulation. The GPCR polypeptide and nucleic acid are also useful for
 CC diagnosing the presence of or predisposition to a disease associated
 CC with altered levels of GPCR, particularly cancer. The GPCR nucleic acid
 CC and polypeptide are especially useful in therapeutic or prophylactic

CC applications for disorders of the neuro-olfactory system, e.g. those
 CC induced by trauma, surgery and/or neoplastic disorders. The DNA encoding
 CC the protein is useful in gene therapy for treating the above conditions.
 CC Furthermore, the nucleic acids and polypeptides are useful in treating
 CC adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune
 CC response, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's
 CC disease, multiple sclerosis or Albright hereditary osteodystrophy. These
 CC are also useful in developing powerful assay system for functional
 CC analysis of various human disorders, as well as in diagnostic
 CC applications. The present amino acid sequence represents one of a
 CC collection (AA068134-AA068153) of human GPCR proteins of the invention.
 CC
 XX
 SO Sequence 315 AA:
 QY 1 MGRVNVNSTYDGFLLGIFSHSOTDLVLFSAVMVVFVALCGNVLLFLIYLDAGLHTPM 60
 DB 1 MGRVNVNSTYDGFLLGIFSHSOTDLVLFSAVMVVFVALCGNVLLFLIYLDAGLHTPM 60
 QY 61 YFELSQSLMDMLVCNIVPKMANFLSGRKSISFEVCGIQIGFVSLVSGELLLGLMA 120
 DB 61 YFELSQSLMDMLVCNIVPKMANFLSGRKSISFEVCGIQIGFVSLVSGELLLGLMA 120
 QY 121 YDHVAVSHPLHPILMNQVCIQTGSSWAFGIDGVYOMVAMGLPYCGSSVDHFFC 180
 DB 121 YDHVAVSHPLHPILMNQVCIQTGSSWAFGIDGVYOMVAMGLPYCGSSVDHFFC 180
 QY 181 EVOALLKACADTSLFDTLLFACCVFMLLPFSIIMASYACIIGAVLRISAOAMKKALA 240
 DB 181 EVOALLKACADTSLFDTLLFACCVFMLLPFSIIMASYACIIGAVLRISAOAMKKALA 240
 QY 241 TCSSHLTAVTLEFYGAAMFMYLRPRRYRAPSHDKVASIFTYVLPMLNPLIYSLRNGEVMG 300
 DB 241 TCSSHLTAVTLEFYGAAMFMYLRPRRYRAPSHDKVASIFTYVLPMLNPLIYSLRNGEVMG 300
 QY 301 ALRKGDLRCRIGSOH 315
 DB 301 ALRKGDLRCRIGSOH 315
 RESULT 4
 AAG71699
 ID AAG71699 standard; Protein; 315 AA.
 XX
 AC AAG71699;
 XX
 DT 30-JUL-2001 (first entry)
 DE Human olfactory receptor polypeptide, SEQ ID NO: 1380.
 XX
 KW Human: Olfactory receptor; OR; primary scent determination;
 KW secondary scent determination; polypeptide library; odour receptor;
 KW scent profile; scent fingerprint; scent representation.
 XX
 OS Homo sapiens.
 XX
 PN WO200127158-A2.
 PD 19-APR-2001.
 XX
 PF 06-OCT-2000; 2000MO-US27582.
 XX
 PR 08-OCT-1999; 99US-0158615.
 PR 24-FEB-2000; 2000US-0184809.
 XX
 PA (DIGI-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX
 PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 XX

DR WPI: 2001-290713/30.
 XX New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists -
 XX
 PS Claim 11: Page 860-861, 1857pp: English.
 CC
 CC The present sequence is an olfactory receptor which is encoded by
 CC one of a number of novel polynucleotides. The polynucleotides can be
 CC used in screening for olfactory agonists and antagonists. The methods
 CC allow for the determination of primary scents and the identification
 CC of the odour receptors used to detect these primary scents. The methods
 CC also enable determination of secondary scents and the identification of
 CC combinations of odour receptors that are involved in detecting such
 CC secondary scents. This enables the construction of a scent representation
 CC (also called a scent fingerprint or scent profile), which may be used to
 CC re-create and edit scents. Libraries of olfactory receptors are useful
 CC for determining the interaction pattern of a composition with the
 CC receptors, and can be used for determining differences in the olfactory
 CC faculties of different individuals.
 CC
 XX Sequence 315 AA;
 SQ
 Query Match 87.6%; Score 1432; DB 22; Length 315;
 Best Local Similarity 87.0%; Pred. No. 1.4e-155;
 Matches 274; Conservative 13; Mismatches 28; Indels 0; Gaps 0;
 QY 1 MGRWNOSYTDGFFLLGIFSHSOTDVLFSAVMVFYALCGNVLLIFLYLDAGLHTPM 60
 DB 1 METWVNOSTDGFLLGIFSHSTADLVFSVMAVFYALCGNVLLIFLYLDAGLHTPM 60
 QY 61 YFFLSQSLMDLMLVCNVPKMANFLSGRKSISFVGGGIDIGFVSLVSGEGLLGLMA 120
 DB 61 YFFLSQSLMDLMLVCNVPKMANFLSGRKSISFVGGGIDIGFVSLVSGEGLLGLMA 120
 QY 121 YDHYVAVSHPLHYPLIMNORVCLQITGSSMAFGIIDGLIQWVVMNPFYCGLRKVNHEFC 180
 DB 121 YDHYVAVSHPLHYPLIMNORVCLQITGSSMAFGIIDGLIQWVVMNPFYCGLRKVNHEFC 180
 QY 181 EVOALLKACADTSLFDTLFACCVFMLLPFSITIMASYACILCAVLRINSAQAMKALA 240
 DB 181 EMLSLKLACVDTSLFEKVFACCVFMLLPFSITIVASYAHILCTVLOMHSQAQMKALA 240
 QY 241 TCSSHLRAVTLFYGAAMFYLRPRRYRAPSHDKVASIFYVLTPLMLPLIYSLNREVMG 300
 DB 241 TCSSHLRAVTLFYGAAMFYLRPRRYRAPSHDKVASIFYVLTPLMLPLIYSLNREVMG 300
 QY 301 ALRKGDLRCRIGSOH 315
 DB 301 ALRKGDLRCRIGSOH 315
 RESULT 5
 AAG71792
 ID AAG71792 standard; Protein: 315 AA.
 AC AAG71792;
 DT 30-JUL-2001 (first entry)
 DE Human olfactory receptor polypeptide, SEQ ID NO: 1473.
 XX
 XX Human: olfactory receptor; OR: primary scent determination;
 KM secondary scent determination; polypeptide library; odour receptor;
 KW scent profile; scent fingerprint; scent representation.
 XX
 OS Homo sapiens.
 XX
 XX WO200127158-A2.
 PN 19-APR-2001.
 PD 06-OCT-2000; 2000MO-US27582.
 PF

XX 08-OCT-1999; 99US-0158615.
 PR 24-FEB-2000; 2000US-0184809.
 XX
 XX (DIGI-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 PA
 PI Bellenson J, Smith D, Lancel D, Glusman G, Fuchs T, Yanai I;
 XX WPI: 2001-290713/30.
 DR
 XX New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists -
 XX
 PS Claim 11: Page 932; 1857pp: English.
 CC
 CC The present sequence is an olfactory receptor which is encoded by
 CC one of a number of novel polynucleotides. The polynucleotides can be
 CC used in screening for olfactory agonists and antagonists. The methods
 CC allow for the determination of primary scents and the identification
 CC of the odour receptors used to detect these primary scents. The methods
 CC also enable determination of secondary scents and the identification of
 CC combinations of odour receptors that are involved in detecting such
 CC secondary scents. This enables the construction of a scent representation
 CC (also called a scent fingerprint or scent profile), which may be used to
 CC re-create and edit scents. Libraries of olfactory receptors are useful
 CC for determining the interaction pattern of a composition with the
 CC receptors, and can be used for determining differences in the olfactory
 CC faculties of different individuals.
 CC
 XX Sequence 315 AA;
 SQ
 Query Match 87.6%; Score 1432; DB 22; Length 315;
 Best Local Similarity 87.0%; Pred. No. 1.4e-155;
 Matches 274; Conservative 13; Mismatches 28; Indels 0; Gaps 0;
 QY 1 MGRWNOSYTDGFFLLGIFSHSOTDVLFSAVMVFYALCGNVLLIFLYLDAGLHTPM 60
 DB 1 METWVNOSTDGFLLGIFSHSTADLVFSVMAVFYALCGNVLLIFLYLDAGLHTPM 60
 QY 61 YFFLSQSLMDLMLVCNVPKMANFLSGRKSISFVGGGIDIGFVSLVSGEGLLGLMA 120
 DB 61 YFFLSQSLMDLMLVCNVPKMANFLSGRKSISFVGGGIDIGFVSLVSGEGLLGLMA 120
 QY 121 YDHYVAVSHPLHYPLIMNORVCLQITGSSMAFGIIDGLIQWVVMNPFYCGLRKVNHEFC 180
 DB 121 YDHYVAVSHPLHYPLIMNORVCLQITGSSMAFGIIDGLIQWVVMNPFYCGLRKVNHEFC 180
 QY 181 EVOALLKACADTSLFDTLFACCVFMLLPFSITIMASYACILCAVLRINSAQAMKALA 240
 DB 181 EMLSLKLACVDTSLFEKVFACCVFMLLPFSITIVASYAHILCTVLOMHSQAQMKALA 240
 QY 241 TCSSHLRAVTLFYGAAMFYLRPRRYRAPSHDKVASIFYVLTPLMLPLIYSLNREVMG 300
 DB 241 TCSSHLRAVTLFYGAAMFYLRPRRYRAPSHDKVASIFYVLTPLMLPLIYSLNREVMG 300
 QY 301 ALRKGDLRCRIGSOH 315
 DB 301 ALRKGDLRCRIGSOH 315
 RESULT 6
 AAG72461
 ID AAG72461 standard; Protein: 315 AA.
 AC AAG72461;
 DT 30-JUL-2001 (first entry)
 DE Human OR-like polypeptide query sequence, SEQ ID NO: 2142.
 XX
 XX Human: olfactory receptor; OR: primary scent determination;
 KM secondary scent determination; polypeptide library; odour receptor;
 KW


```

Db      1  METVWNSYTDGFFLLGIFSHSTADLVLSVMAVFVALCGNVLLIFLLYMDPHLHTPM 60
QY      61  YFPLSOLSLMDLMLVCNTVPMKMANFLSGRKSISFVCGGIDIGFVSLVSGEGLLGLMA 120
Db      61  YFPLSOLSLMDLMLVCNTVPMKMANFLSGRKSISFVCGGIDIGFVSLVSGEGLLGLMA 120
QY      121  YDHYVAVSHPLHYPLIMNORVCLQITGSSNAFGIIDDVIOVMAMGLPYCGSRSDVHFFC 180
Db      121  YDHYVAVSHPLHYPLIMNORVCLQITGSSNAFGIIDDVIOVMAMGLPYCGSRSDVHFFC 180
QY      181  EVOALLKLACADTSIFPTLFLFACCVFMLLPFSITMASVACIIIGAVLRISQAAMKKALA 240
Db      181  EVOALLKLACADTSIFPTLFLFACCVFMLLPFSITMASVACIIIGAVLRISQAAMKKALA 240
QY      241  TCSSHLTAVTLEFYGAAMFYLRPRRYRAPSHDKVASIFTYVLTPMLNPLIYSLNNGEVMG 300
Db      241  TCSSHLTAVTLEFYGAAMFYLRPRRYRAPSHDKVASIFTYVLTPMLNPLIYSLNNGEVMG 300
QY      301  ALRKGIDRCRIGSOH 315
Db      301  ALRKGIDRCRIGSOH 315

RESULT 8
AAE18656 standard; Protein: 315 AA.
XX
AC      AAE18656:
XX
DT      17-MAY-2002 (first entry)
XX
DE      Human G-protein coupled receptor (GCREC-17).
XX
KW      Human: G-protein coupled receptor; GCREC-17; cell proliferative disorder;
KW      neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;
KW      metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;
KW      Pick's disease; Huntington's disease; Parkinson's disease; hypertension;
KW      atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;
KW      osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW      anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;
KW      transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;
KW      neurotropic; neuroprotective; cardiant; immunosuppressive; anorectic;
KW      virucide; receptor.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      Peptide    1..24
FT      Protein    25..315
FT      Domain     /label= Human_mature_GCREC-17
FT      Domain     26..46
FT      Domain     /note= "Transmembrane domain"
FT      Domain     98..119
FT      Domain     /note= "Transmembrane domain"
FT      Domain     201..220
FT      Domain     /note= "Transmembrane domain"
XX
PN      WO200210387-A2.
XX
PD      07-FEB-2002.
XX
PF      25-JUL-2001; 2001WO-US23433.
XX
PR      27-JUL-2000; 2000US-221478P.
PR      03-AUG-2000; 2000US-223268P.
PR      21-AUG-2000; 2000US-227054P.
PR      08-SEP-2000; 2000US-231121P.
PR      13-SEP-2000; 2000US-232243P.
PR      15-SEP-2000; 2000US-232691P.
PR      22-SEP-2000; 2000US-235146P.
XX

```

```

PA      (INCYTE GENOMICS INC.
XX      Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR.
PI      Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Walla NK.
PI      Hafalia AJA, Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L.
PI      Graul RC, Warren BA, Lee EA, Ding L.
XX
DR      WPI, 2002-188744/24.
XX      N-PSDB; AAD29683.
XX
PT      New human G-protein coupled receptor polypeptide for diagnosis,
PT      prevention and treatment of cell proliferative, neurological,
PT      cardiovascular, gastrointestinal, autoimmune/inflammatory, and
PT      metabolic disorders.
XX
PS      Claim 1; Page 135-136; 150pp; English.
XX
CC      The invention relates to novel human G-protein coupled receptors
CC      (GCREC) and their encoding polynucleotides. GCREC is useful as an
CC      immunogen for preparing monoclonal and polyclonal antibodies. GCREC is
CC      useful for diagnosing, treating and preventing a cell proliferative
CC      disorder (e.g., hepatitis, psoriasis, cancer), a neurological disorder
CC      (e.g., epilepsy, Alzheimer's disease, Pick's disease, Huntington's
CC      disease, Parkinson's disease), a cardiovascular disorder (e.g.,
CC      atherosclerosis, hypertension, myocardial infarction), gastrointestinal
CC      disorder (e.g., gastritis, cirrhosis, Crohn's disease), an autoimmune/
CC      inflammatory disorder (e.g., acquired immunodeficiency syndrome (AIDS),
CC      allergy, anaemia, asthma, rheumatoid arthritis), a metabolic disorder
CC      (e.g., diabetes, obesity, osteoporosis), and viral infections. GCREC is
CC      useful in a number of drug screening techniques, and to analyse the
CC      proteome of a tissue or cell type. GCREC is useful for creating knockin
CC      humanised animals or transgenic animals to model human diseases, in
CC      somatic or germ-line gene therapy, to generate a transcript image of a
CC      tissue or cell type, for detecting differences in the chromosomal
CC      location due to translocation, inversion, etc., among normal, carrier
CC      or affected individuals, and as hybridization probes for mapping
CC      naturally occurring genomic sequences. GCREC is useful in Southern or
CC      Northern analysis, dot blot or other membrane-based technologies, in PCR
CC      technologies, in dipstick, pin, multiformat enzyme linked immunosorbent
CC      (ELISA)-like assays, and in microarrays utilising fused or tissues from
CC      human GCREC-17.
XX
SQ      Sequence 315 AA:
XX
Query Match      87.6%; Score 1432; DB 23; Length 315;
Best Local Similarity      87.0%; Pred. No. 1.4e-155;
Matches 274; Conservative 13; Mismatches 28; Indels 0; Gaps 0;
QY      1  MGRWNSYTDGFFLLGIFSHSTADLVLSVMAVFVALCGNVLLIFLLYMDAGLHTPM 60
Db      1  METVWNSYTDGFFLLGIFSHSTADLVLSVMAVFVALCGNVLLIFLLYMDPHLHTPM 60
QY      61  YFPLSOLSLMDLMLVCNTVPMKMANFLSGRKSISFVCGGIDIGFVSLVSGEGLLGLMA 120
Db      61  YFPLSOLSLMDLMLVCNTVPMKMANFLSGRKSISFVCGGIDIGFVSLVSGEGLLGLMA 120
QY      121  YDHYVAVSHPLHYPLIMNORVCLQITGSSNAFGIIDDVIOVMAMGLPYCGSRSDVHFFC 180
Db      121  YDHYVAVSHPLHYPLIMNORVCLQITGSSNAFGIIDDVIOVMAMGLPYCGSRSDVHFFC 180
QY      181  EVOALLKLACADTSIFPTLFLFACCVFMLLPFSITMASVACIIIGAVLRISQAAMKKALA 240
Db      181  EVOALLKLACADTSIFPTLFLFACCVFMLLPFSITMASVACIIIGAVLRISQAAMKKALA 240
QY      241  TCSSHLTAVTLEFYGAAMFYLRPRRYRAPSHDKVASIFTYVLTPMLNPLIYSLNNGEVMG 300
Db      241  TCSSHLTAVTLEFYGAAMFYLRPRRYRAPSHDKVASIFTYVLTPMLNPLIYSLNNGEVMG 300
QY      301  ALRKGIDRCRIGSOH 315
Db      301  ALRKGIDRCRIGSOH 315

```

```

RESULT 9
AA085223
ID AA085223 standard; Protein; 315 AA.
XX
AC AA085223;
XX
DT 08-MAY-2002 (first entry)
XX
DE G-coupled olfactory receptor #84.
XX
KW Human; olfactory G-coupled receptor; sensory perception of odourant;
KM odour composition; taste composition.
XX
OS Homo sapiens.
XX
PN M0200198526-A2.
XX
PD 27-DEC-2001.
XX
PF 22-JUN-2001; 2001MO-US20122.
XX
PR 22-JUN-2000; 2000US-213812P.
PR 13-MAR-2001; 2001US-0804291.
XX
PA (SENO-) SENOMYX INC.
XX
PI Zozulya S, Stryer L;
XX
DR WPI; 2002-083330/11.
DR N-PSDB; ABK37582.
XX
PT Representing sensory perception of one or more odourants for the
PT identification and design of tastes and odours comprises providing a
PT representative group of n olfactory receptors -
XX
Claim 1; Page 96; 182pp; English.
XX
XX
XX The invention relates to a method of representing sensory perception of
XX one or more odourants. The method comprises: (a) providing a
XX representative class of n olfactory receptors or ligand binding domains
XX (LBD) of these receptors; (b) measuring values X1 to Xn representative
XX of at least one activity of one or more odourants selected from:
XX (i) binding one or more odourants to the LBD of at least one of the n
XX olfactory receptors; (ii) activating at least one of the n
XX olfactory receptors with the one or more odourants; and (iii) blocking at
XX least one of the n olfactory receptors with the one or more odourants;
XX and (c) generating a representation of sensory perception from the values
XX X1 to Xn. The representation of the sensory perception of odourants is
XX useful for the design and formulation of odour and taste compositions.
XX CC AA085140-AA085393 represent human olfactory G-coupled receptor amino acid
XX sequences of the invention.
XX
SQ Sequence 315 AA:
XX
Query Match 87.6%; Score 1432; DB 23; Length 315;
Best Local Similarity 87.0%; Pred. No. 1.4e-155;
Matches 274; Conservative 13; Mismatches 28; Indels 0; Gaps 0;
XX
QY 1 MGRWVNSYTDGFFLLGIFSHSOTDLVLSAVVVFVLCGAVLLFLIYDAGLHTPM 60
DB 1 MGRWVNSYTDGFFLLGIFSHSOTDLVLSAVVVFVLCGAVLLFLIYDAGLHTPM 60
XX
QY 61 YFFLSQSLMDLMLVNCINIVKMANFLSGRKSISFVCGGIQIGFPIVSGSEGLLGIMA 120
DB 61 YFFLSQSLMDLMLVNCINIVKMANFLSGRKSISFVCGGIQIGFPIVSGSEGLLGIMA 120
XX
QY 121 YDHYVAVSHPLATPIIMNQRYCLQITGSSWAFGIIDGVLIQMAVAMGIPYCGSRSDVDFEC 180
DB 121 YDHYVAVSHPLATPIIMNQRYCLQITGSSWAFGIIDGVLIQMAVAMGIPYCGSRSDVDFEC 180
XX
QY 181 EYVALKLACADTSLFDTLFACCVFMLLPSTITMASVACIIIGAVIRLSAQAAMKALA 240
DB 181 EYVALKLACADTSLFDTLFACCVFMLLPSTITMASVACIIIGAVIRLSAQAAMKALA 240
XX

```

```

DB 181 EMLSLKLACVDTSLEFEKVIACCVFMLLPFSIIIVASVAHIIIGTVLQMSAQAAMKALA 240
QY 241 TCSSHLLAATVLTGCAAMFVLPRPRAPSHDKVASTFYVLPMLNPLIYSJLNGEVMG 300
DB 241 TCSSHLLAATVLTGCAAMFVLPRPRAPSHDKVASTFYVLPMLNPLIYSJLNGEVMG 300
QY 301 ALRKGLDRCRIGSQH 315
DB 301 ALRKGLDRCRIGSQH 315
XX
RESULT 10
AAE16175
ID AAE16175 standard; Protein; 271 AA.
XX
AC AAE16175;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human G-protein coupled receptor 6 (GPRC-6) protein.
XX
KW Human; G-protein coupled receptor 6; cell proliferative disorder;
KW arteriosclerosis; hepatitis; cancer; neurological disorder; epilepsy;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW atherosclerosis; hypertension; myocardial infarction; peptic ulcer;
KW gastrointestinal disorder; dysphagia; anorexia; autoimmune disorder;
KW acquired immune deficiency syndrome; inflammatory disorder; infection;
KW Addison's disease; allergy; Grave's disease; metabolic disorder; AIDS;
KW diabetes; obesity; osteoporosis; gene therapy; GPRC-6.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 26..46 /note="Transmembrane domain"
FT Domain 42..155 /note="7 transmembrane receptor domain"
FT Domain 111..127 /note="G-protein coupled receptor motif"
FT Domain 223..247 /note="7 transmembrane receptor domain"
XX
W0200187937-A2.
XX
22-NOV-2001.
XX
17-MAY-2001; 2001MO-US16285.
XX
18-MAY-2000; 2000US-205628P.
XX
22-MAY-2000; 2000US-206222P.
XX
25-MAY-2000; 2000US-207566P.
XX
02-JUN-2000; 2000US-208834P.
XX
02-JUN-2000; 2000US-208861P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Patterson C, Lu DAM, Thornton M, Lu Y, Tribouley CM, Graul R,
PI Khan FA, Gandhi AR, Malia NK, Nguyen DA, Yue H, Hafajia A;
PI Elliott VS, Lai P, Reddy R, Kallick DA, Tang TT, Au-Young J;
XX
WPI; 2002-089844/12.
XX
N-PSDB; AAD26374.
XX
PT Novel G-protein coupled receptors and polynucleotides useful for
PT diagnosis, treatment and prevention of disorders of cell proliferation,
PT neurological, cardiovascular, metabolic disorders and viral infections
XX
Claim 1; Page 108; 115pp; English.
XX
XX The invention relates to human G-protein coupled receptor (GPRC)
XX polypeptides and polynucleotides. GPRC polypeptides are useful for
XX screening compounds that modulate their activity. They are useful in
XX

```


XX 18-SEP-2002 (first entry)
DT
XX
DE Human GPCR protein #4.
XX
KW Human: G-protein coupled receptor; GPCR; cardiact; antiatherosclerotic;
KW antiatherosclerotic; hypotensive; anorectic; cytosolic; immunosuppressive;
KW neuroprotective; norepinephrine; antiparkinsonian; antidiabetic; gene therapy;
KW cardiomyopathy; atherosclerosis; cell signal processing; diabetes;
KW metabolic pathway; hypertension; obesity; cancer; Alzheimer's disease;
KW graft versus host disease; Parkinson's disorder; dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO20025557-A2.
XX
PD 18-JUL-2002.
XX
PF 18-DEC-2001; 2001WO-US49111.
XX
PR 18-DEC-2000; 2000US-256635P.
PR 21-DEC-2000; 2000US-257876P.
PR 04-JAN-2001; 2001US-259743P.
PR 10-JAN-2001; 2001US-260718P.
PR 12-JAN-2001; 2001US-261498P.
PR 24-JAN-2001; 2001US-263689P.
PR 08-FEB-2001; 2001US-267464P.
PR 22-FEB-2001; 2001US-271021P.
PR 14-MAR-2001; 2001US-275946P.
PR 23-MAR-2001; 2001US-278150P.
PR 16-AUG-2001; 2001US-312902P.
XX
XX
XX (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Kekuda R, Colman SD, Spytek KA, Ballinger RA;
PI Verneet CAM, Li L, Shenoy S, Casman SJ, Gusev V;
XX
DR WPI: 2002-500839/53.
DR N-PSDB; ABQ88046.
XX
XX New G-protein coupled receptor polypeptides, encoding nucleic acids and
PT immunospecific antibodies, useful for treating cardiomyopathy,
PT atherosclerosis, diabetes or disorder related to cell signal processing
PT and metabolic pathway modulation -
XX
PS Claim 1; Page 86; 146pp; English.
XX
XX The invention relates to a novel nucleic acid molecule comprising a
CC nucleic acid sequence encoding a G-protein coupled receptor (GPCR)
CC polypeptide. The proteins of the invention have cardiac,
CC antiatherosclerotic, antidiabetic, hypotensive, anorectic, cytosolic,
CC immunosuppressive, neuroprotective, norepinephrine, antiparkinsonian, and
CC antipneumatic activity. The nucleic acids of the invention may have a use
CC in gene therapy. The sequences of the invention are useful for treating
CC or preventing a GPCR-associated disorder, preferably cardiomyopathy,
CC atherosclerosis or a disorder related to cell signal processing and
CC metabolic pathway modulation. An antibody of the invention is useful for
CC treating diabetes or a disorder related to cell signal processing and
CC metabolic pathway modulation. GPCR-associated disorders include
CC hypertension, obesity, cancer, graft versus host disease, Alzheimer's
CC disease, Parkinson's disorder and dyslipidaemia. The sequences shown in
CC ABQ88043-ABQ88093 encode the GPCR proteins of the invention, shown in
CC ABP61129-ABP61179.
XX
XX
SO Sequence 263 AA:
Query Match 70.1%; Score 1146; DB 23; Length 263;
Best Local Similarity 99.5%; Pred. No. 7.3e-123;
Matches 221; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 61 YFFLSQSLMDLMLVCNIVPKMANFLSGRKSISFVGGCIGDGFVSLVSGEGLLGMA 120
DB 61 YFFLSQSLMDLMLVCNIVPKMANFLSGRKSISFVGGCIGDGFVSLVSGEGLLGMA 120
QY 121 YDRVAVNSHPPLHYPIILMNQRCLOITGSSNAFGIIDDYIOWAAMGLPYGSRSDHFFC 180
DB 121 YDRVAVNSHPPLHYPIILMNQRCLOITGSSNAFGIIDDYIOWAAMGLPYGSRSDHFFC 180
QY 181 EVOALLKACADTSLFPTLFFACCVFMLLPFSITMASYACT 222
DB 181 EVOALLKACADTSLFPTLFFACCVFMLLPFSITMASYACT 222
RESULT 13
AA96666
ID AA96666 standard; Protein: 223 AA.
XX
XX AA96666;
XX
DT 26-SEP-2000 (first entry)
XX
DE Murine olfactory receptor ligand-binding region 5.
XX
KW Degenerate primer; ligand-binding region; olfactory receptor;
KW Transmembrane domain; N-terminal; plasma membrane; translocation domain;
KW human rhodopsin receptor; odorant; toxicity; olfactory response.
XX
OS Mus musculus.
XX
PN WO200035274-A1.
XX
PD 22-JUN-2000.
XX
PF 17-DEC-1999; 99WO-US30221.
XX
PR 17-DEC-1998; 98US-0112605.
XX
PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Reed RR, Krautwurst D, Yau KW;
XX
DR WPI: 2000-431471/37.
DR N-PSDB; AA51179.
XX
XX Primer pair for amplifying olfactory receptor nucleic acid, useful e.g.
PT for producing receptor libraries used in e.g. screening odorants for
PT toxicity
XX
PS Claim 18; Page 32; 61pp; English.
XX
CC AA96662-80 are murine odorant/ligand binding regions of olfactory
CC receptors. The coding sequences were amplified using degenerate primers
CC shown in AA51163-64.
CC The primers can be used to generate a library of OR-IBR comprising the
CC transmembrane domains (TMD) II-VI, III-VI, II-VII or III-VII. Chimeric
CC nucleic acids encoding a 7-TMD protein comprise a sequence encoding an
CC N-terminal plasma membrane translocation domain (especially the sequence
CC shown in AA96658), a first TMD and the OR-IBR. The translocation domain
CC was initially derived from the N-terminus of the human rhodopsin
CC receptor. IBR produced by amplification with the primers are useful for
CC generating new odorants, to screen for toxicity or therapeutic activity
CC in odorants, and altering an animal's olfactory response (claimed).
XX
XX
SO Sequence 223 AA:
Query Match 63.9%; Score 1044; DB 21; Length 223;
Best Local Similarity 89.7%; Pred. No. 3e-111;
Matches 200; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

```

OY 125 VAVSHPLHPIILMNQVCIQTGSSNAFGIIDGVIOVMAMGLPYCGSRSDHFECEVOA 184
DB 61 VAIASHLPHPIILMSOKVCIQIAGSSNAFGIIDGVIOVMAMSLPYCGSRIDHFECEVPA 120
OY 185 LKACADTSLFPTLLFACCVFEMLLPFSTIMASYACIIGAVIRISAQAMKALATCSS 244
DB 121 LKACADTSLFPTLLFACCVFEMLLPFSTITVSVARIIIGAVIRHMSAQSRRKALATCSS 180
OY 245 HLTAATLFGAAMFYLRPRRYRAPSHDKVASIFYYVLTPLMLN 287
DB 181 HLTAATLFGAAMFYLRPRRYRAPSHDKVASIFYYVLTPLMLN 223

RESULT 14
AAY96670
ID AAY96670 standard; Protein: 223 AA.
XX
AC AAY96670;
XX
DT 26-SEP-2000 (first entry)
XX
DE Murine olfactory receptor ligand-binding region 9.
XX
KW Degenerate primer; ligand-binding region; olfactory receptor;
KW Transmembrane domain; N-terminal; plasma membrane; translocation domain;
KW human rhodopsin receptor; odorant; toxicity; olfactory response.
XX
OS Mus musculus.
XX
PN WO200035274-A1.
XX
PD 22-JUN-2000.
XX
PF 17-DEC-1999; 99WO-US30221.
XX
PR 17-DEC-1998; 98US-0112605.
XX
PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Reed RR, Krautwurst D, Yau KW;
XX
DR WPI: 2000-431471/37.
XX
DR N-PSDB; AAA51183.
XX
PT Primer pair for amplifying olfactory receptor nucleic acid, useful e.g.
PT for producing receptor libraries used in e.g. screening odorants for
PT toxicity
XX
PS Claim 18; Page 33; 61pp; English.
XX
CC AAY96662-80 are murine odorant/ligand binding regions of olfactory
CC receptors. The coding sequences were amplified using degenerate primers
CC shown in AAA51183-64.
CC The primers can be used to generate a library of OR-LBR comprising the
CC transmembrane domains (TMD) II-VI, III-VI, II-VII or III-VII. Chimeric
CC nucleic acids encoding a 7-TMD protein comprise a sequence encoding an
CC N-terminal plasma membrane translocation domain (especially the sequence
CC shown in AAY96658), a first TMD and the OR-LBR. The translocation domain
CC was initially derived from the N-terminus of the human rhodopsin
CC receptor. LBR produced by amplification with the primers are useful for
CC generating new odorants, to screen for toxicity or therapeutic activity
CC in odorants, and altering an animal's olfactory response (claimed).
XX
SQ Sequence 223 AA;

```

Query Match 63.9%; Score 1044; DB 21; Length 223;
 Best Local Similarity 89.7%; Pred. No. 3e-111;
 Matches 200; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

```

OY 65 SOLSLMDLMLVCNIVPKMAANFLSGRKSISFVGGCIQIGFVSLVSGELLLGLMAVDHY 124
DB 1 SOLSLMDLMLVCNIVPKMAANFLSGRKSISFVGGCIQIGFVSLVSGELLLGLMAVDHY 60

```

```

OY 125 VAVSHPLHPIILMNQVCIQTGSSNAFGIIDGVIOVMAMGLPYCGSRSDHFECEVOA 184
DB 61 VAIASHLPHPIILMSOKVCIQIAGSSNAFGIIDGVIOVMAMSLPYCGSRIDHFECEVPA 120
OY 185 LKACADTSLFPTLLFACCVFEMLLPFSTIMASYACIIGAVIRISAQAMKALATCSS 244
DB 121 LKACADTSLFPTLLFACCVFEMLLPFSTITVSVARIIIGAVIRHMSAQSRRKALATCSS 180
OY 245 HLTAATLFGAAMFYLRPRRYRAPSHDKVASIFYYVLTPLMLN 287
DB 181 HLTAATLFGAAMFYLRPRRYRAPSHDKVASIFYYVLTPLMLN 223

RESULT 15
AAY96668
ID AAY96668 standard; Protein: 223 AA.
XX
AC AAY96668;
XX
DT 26-SEP-2000 (first entry)
XX
DE Murine olfactory receptor ligand-binding region 7.
XX
KW Degenerate primer; ligand-binding region; olfactory receptor;
KW Transmembrane domain; N-terminal; plasma membrane; translocation domain;
KW human rhodopsin receptor; odorant; toxicity; olfactory response.
XX
OS Mus musculus.
XX
PN WO200035274-A1.
XX
PD 22-JUN-2000.
XX
PF 17-DEC-1999; 99WO-US30221.
XX
PR 17-DEC-1998; 98US-0112605.
XX
PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Reed RR, Krautwurst D, Yau KW;
XX
DR WPI: 2000-431471/37.
XX
DR N-PSDB; AAA51181.
XX
PT Primer pair for amplifying olfactory receptor nucleic acid, useful e.g.
PT for producing receptor libraries used in e.g. screening odorants for
PT toxicity
XX
PS Claim 18; Page 32-33; 61pp; English.
XX
CC AAY96662-80 are murine odorant/ligand binding regions of olfactory
CC receptors. The coding sequences were amplified using degenerate primers
CC shown in AAA51183-64.
CC The primers can be used to generate a library of OR-LBR comprising the
CC transmembrane domains (TMD) II-VI, III-VI, II-VII or III-VII. Chimeric
CC nucleic acids encoding a 7-TMD protein comprise a sequence encoding an
CC N-terminal plasma membrane translocation domain (especially the sequence
CC shown in AAY96658), a first TMD and the OR-LBR. The translocation domain
CC was initially derived from the N-terminus of the human rhodopsin
CC receptor. LBR produced by amplification with the primers are useful for
CC generating new odorants, to screen for toxicity or therapeutic activity
CC in odorants, and altering an animal's olfactory response (claimed).
XX
SQ Sequence 223 AA;

```

Query Match 63.6%; Score 1040; DB 21; Length 223;
 Best Local Similarity 89.2%; Pred. No. 8.5e-111;
 Matches 199; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

```

OY 65 SOLSLMDLMLVCNIVPKMAANFLSGRKSISFVGGCIQIGFVSLVSGELLLGLMAVDHY 124
DB 1 SOLSLMDLMLVCNIVPKMAANFLSGRKSISFVGGCIQIGFVSLVSGELLLGLMAVDHY 60

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:59:16 ; Search time 18 Seconds
(without alignments)
1682.352 Million cell updates/sec

Title: US-09-634-109d-2
Perfect score: 1634
Sequence: 1 MGRWVQSYDGFLLGIFS.....GEVMGALRKGLDRICRIGSH 315

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 723 | 44.2 | 316 | 2 | A57069 |
| 2 | 715 | 43.8 | 315 | 2 | JC5836 |
| 3 | 696 | 42.6 | 312 | 2 | A46247 |
| 4 | 654.5 | 40.1 | 312 | 2 | A48413 |
| 5 | 631 | 38.6 | 305 | 2 | S29711 |
| 6 | 629 | 38.5 | 327 | 2 | F23701 |
| 7 | 608 | 37.2 | 309 | 1 | S51356 |
| 8 | 607 | 37.1 | 318 | 2 | JC5202 |
| 9 | 601 | 36.8 | 320 | 2 | S20573 |
| 10 | 595 | 36.4 | 315 | 2 | JC5201 |
| 11 | 591.5 | 36.2 | 313 | 2 | S20571 |
| 12 | 591 | 36.2 | 313 | 2 | B23701 |
| 13 | 588 | 36.0 | 311 | 2 | JC5200 |
| 14 | 583 | 35.7 | 314 | 2 | S29707 |
| 15 | 578 | 35.4 | 319 | 2 | JC5624 |
| 16 | 573 | 35.1 | 304 | 2 | S29709 |
| 17 | 571 | 34.9 | 333 | 2 | A23701 |
| 18 | 570 | 34.9 | 314 | 2 | S20572 |
| 19 | 570 | 34.9 | 314 | 2 | A37286 |
| 20 | 568 | 34.8 | 314 | 2 | H23701 |
| 21 | 562 | 34.4 | 307 | 2 | S29710 |
| 22 | 560 | 34.3 | 312 | 2 | S29708 |
| 23 | 558 | 34.1 | 310 | 2 | E23701 |
| 24 | 552 | 33.8 | 315 | 2 | JC4658 |
| 25 | 545.5 | 33.4 | 317 | 2 | D23701 |
| 26 | 533 | 32.6 | 222 | 2 | D40745 |
| 27 | 531 | 32.5 | 312 | 2 | I23701 |
| 28 | 522 | 31.9 | 312 | 2 | G23701 |
| 29 | 516 | 31.6 | 157 | 2 | S58073 |

| | | | | | | |
|----|-------|------|-----|---|--------|--------------------|
| 30 | 501.5 | 30.7 | 312 | 2 | A46750 | olfactory receptor |
| 31 | 490 | 30.0 | 311 | 2 | C23701 | olfactory receptor |
| 32 | 483 | 29.6 | 157 | 2 | S58037 | probable olfactory |
| 33 | 481 | 29.4 | 157 | 2 | S58029 | probable olfactory |
| 34 | 479 | 29.3 | 222 | 2 | B40745 | odorant receptor (|
| 35 | 471 | 28.8 | 264 | 2 | PC4369 | olfactory receptor |
| 36 | 441 | 27.0 | 157 | 2 | S58033 | probable olfactory |
| 37 | 437 | 26.7 | 157 | 2 | S58019 | probable olfactory |
| 38 | 424 | 25.9 | 234 | 2 | S29000 | G protein-coupled |
| 39 | 422 | 25.8 | 157 | 2 | S58011 | probable olfactory |
| 40 | 413 | 25.3 | 234 | 2 | S28998 | G protein-coupled |
| 41 | 411.5 | 25.2 | 215 | 2 | I38473 | olfactory receptor |
| 42 | 404 | 24.7 | 216 | 2 | I38480 | olfactory receptor |
| 43 | 397.5 | 24.3 | 232 | 2 | S29001 | G protein-coupled |
| 44 | 396.5 | 24.3 | 328 | 2 | G45774 | odorant receptor 2 |
| 45 | 396 | 24.2 | 216 | 2 | I38481 | olfactory receptor |

ALIGNMENTS

RESULT 1

A57069
olfactory receptor FAT11 - human
C:Species: Homo sapiens (man)
C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 26-Aug-1999
C:Accession: A57069
R:Fan, W., Liu, Y.C., Parlimoo, S., Weissman, S.M.
Genomics 27, 119-123, 1995
A>Title: Olfactory receptor-like genes are located in the human major histocompatibility
A:Reference number: A57069; MUID:95394447; PMID:765158
A:Accession: A57069
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-316 <FAN>
C:Cross-references: GB:I35475; NID:q1041044; PIDN:AB36567.1; PID:9601919
C:Genetics:
A:Gene: GDB:FAT11; OLFR2
A:Cross-references: GDB:I33249; OMIM:600578
A:Map position: 6p21.3-6p21.3
C:Superfamily: olfactory receptor OR14

| Query Match | Score | DB 2 | Length |
|---|-------|---|--------|
| Best local similarity | 46.6% | Pred. No. 3e-54; | |
| Matches 142; Conservative 52; Mismatches 111; Indels 0; Gaps 0; | | | |
| QY | 6 | NQSYTDEFFLLGIFSHSQDVLVPSAVMVVTVALCGNVLLIFLIYDAGLHPMYFELS 65 | |
| DB | 3 | NOSTEGFLLGFSEHPGLGRTLFVDVITSYLLTLVGNLIIILSALDTKLHSPMYFELS 62 | |
| QY | 66 | QLSLMDLMLVCNIVPKMAANFLSGRKSISFVGGIOGFVSVLWGSRGILLGMAVDHY 125 | |
| DB | 63 | NLSFLDLCFTTSCVPQWLAMLWGPCKTISFLDCSVQIFLISGTECIIMKMAWDRIY 122 | |
| QY | 126 | AVSHPLHYPLIMQRYCLOITGSSMAFGIIGVYQWAAAGLPGYCSRSVDHFECEYOAL 185 | |
| DB | 123 | AVQCPHYATIIHPRICMQLASVAVWIGLVSVQRPSTLHLPEFCDDROVDVCEPAL 182 | |
| QY | 186 | LKLAACDTSLEFTLLFACCVEMLLPESITMASYACILGAVLIRSNQAANKKALATCSSH 245 | |
| DB | 183 | IRLSCEDTSYNEIOVAVSVFLLVPLSLVSYGATWAVLRINSATAWRKAFGTCSSH 242 | |
| QY | 246 | LTVVTEFYGAMFMVLRPRRYRAPSHDKVASITVYTLPLNLNLISLNGEVMGALRGC 305 | |
| DB | 243 | LTVVTEFYSSVLAIVYQPKPYAQGRKFEGLFYAVGTDSLNLVYTLRKKEIKRALRRL 302 | |
| QY | 306 | LDRCR 310 | |
| DB | 303 | LKKER 307 | |
| RESULT 2 | | | |
| JC5836 | | | |

RESULT 8

JC5202

Chemoreceptor TB641 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-Aug-1999

C:Accession: JC5202; PC4304

R:Thomas, M.B.; Haines, S.L.; Akesson, R.A.

A:Title: Chemoreceptors expressed in taste, olfactory and male reproductive tissues.

A:Reference number: JC5200; MUID:97080538; PMID:8921863

A:Accession: JC5202

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-318 <THO1>

A:Cross-references: GB:050949; NID:g1256392; PIDN:AAC52911.1; PID:g1256393

A:Accession: PC4304

A:Status: preliminary

A:Molecule type: protein

A:Residues: 145-153;245-253 <THO2>

A:Experimental source: taste bud

C:Comment: This protein is coupled to a GTP-binding protein-mediated signal transduction

C:Gene: tb641

C:Superfamily: olfactory receptor OR14

C:Keywords: olfaction; taste bud; transmembrane protein

F:30-53/Domain: transmembrane #status predicted <TM1>

F:63-84/Domain: transmembrane #status predicted <TM2>

F:106-125/Domain: transmembrane #status predicted <TM3>

F:145-169/Domain: transmembrane #status predicted <TM4>

F:202-224/Domain: transmembrane #status predicted <TM5>

F:243-265/Domain: transmembrane #status predicted <TM6>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

F:276-296/Domain: transmembrane #status predicted <TM7>

A:Residues: 1-320 <PAR>

A:Cross-references: EMBL:X64995; NID:g32092; PIDN:CAA46128.1; PID:g32093

C:Superfamily: olfactory receptor OR14

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match

Best Local Similarity 36.8%; Score 601; DB 2; Length 320;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Matches 117; Conservative 71; Mismatches 112; Indels 16; Gaps 4;

Db 5 NATEVDFEYLLGFGVQNTQCVLEIFVEFVIYVTSMVGNMTGILLINTNSRLQTPMYEFLQ 64
 QY 66 QLSIMDLMLVNCNIPKMAANFLSGRKSISFVCGIOIGFVSLVSGREGLLGLMAVDHYV 125
 Db 65 NLAEVDICYSATIPKMLQSFWEDECSISTYGCYIQLVYTFEATSDCYLLAAVMAVDHYV 124
 QY 126 AVSHPLHYPIILMNQVCLQITGSSMAFGIIDGVIQMWAAAMGLPYCGSRSDHFECEVOAL 185
 Db 125 AICPRLRYPIILMSQVCLLVALSYLMSINSVHTGFTFSLSCNSKNINHFECDDVPI 184
 QY 186 LKACADTSLFDTLIFACCVFMLLPFSIIMASTACILGAVLRIRSAQAMKALATCSSH 245
 Db 185 ISLSCSNNDIMIMLIVFVGFNLFTFVLIFFSYIYIMAILKMSSTAGRKFTSTCASH 244
 QY 246 LTAATLFFGAMFMYLPRRYRAPSHDKVASIFETVLTPLMPLIYSLRNREV 298
 Db 245 LTAATIRGTISTMILQPHSDNSEMKAVAFVGIYVIMPLNPLIYSLRNEV 297

RESULT 11

S20571

olfactory receptor - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 26-Aug-1999

C:Accession: S20571

R:Parmentier, M.; Libert, F.; Schurmans, S.; Schiffmann, S.; Lefort, A.; Eggerickx, D.;

Nature 355, 453-455, 1992

A:Title: Expression of members of the putative olfactory receptor gene family in mammal

A:Reference number: S20571; MUID:92131132; PMID:1370859

A:Accession: S20571

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-313 <PAR>

A:Cross-references: EMBL:X6496; NID:9890; PIDN:CAA46129.1; PID:9891

C:Superfamily: olfactory receptor OR14

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 36.2%; Score 591.5; DB 2; Length 313;

Best Local Similarity 39.7%; Pred. No. 4.8e-43;

Matches 122; Conservative 62; Mismatches 120; Indels 3; Gaps 3;

QY 6 NOSTYDGEFLIGT-FSHQTDLVLFSAVNVYVALCGNVLLIFLIYDAGLHTPMYFEL 64
 Db 5 NOSTYDGEFLIGT-FSHQTDLVLFSAVNVYVALCGNVLLIFLIYDAGLHTPMYFEL 63
 QY 65 SOLSLMDLMLVNCNIPKMAANFLSGRKSISFVCGIOIGFVSLVSGREGLLGLMAVDHY 124
 Db 64 SOLSLMDLMLVNCNIPKMAANFLSGRKSISFVCGIOIGFVSLVSGREGLLGLMAVDHY 123
 QY 125 AVSHPLHYPIILMNQVCLQITGSSMAFGIIDGVIQMWAAAMGLPYCGSRSDHFECEVOA 184
 Db 124 AVSHPLHYPIILMNQVCLQITGSSMAFGIIDGVIQMWAAAMGLPYCGSRSDHFECEVOA 182
 QY 185 LKACADTSLFDTLIFACCVFMLLPFSIIMASTACILGAVLRIRSAQAMKALATCSSH 244
 Db 183 LKACADTSLFDTLIFACCVFMLLPFSIIMASTACILGAVLRIRSAQAMKALATCSSH 242
 QY 245 LTAATLFFGAMFMYLPRRYRAPSHDKVASIFETVLTPLMPLIYSLRNREV 298
 Db 243 LTAATLFFGAMFMYLPRRYRAPSHDKVASIFETVLTPLMPLIYSLRNREV 297
 QY 305 GLDRCRI 311
 Db 303 VICRKKI 309

RESULT 12

B23701

olfactory receptor F5 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 04-Sep-1998

C:Accession: B23701

R:Buck, L.; Axel, R.

Cell 65, 175-187, 1991
 A:Title: A novel multigene family may encode odorant receptors: a molecular basis for
 A:Reference number: A23701; MUID:91191556; PMID:1840504
 A:Accession: B23701
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-313 <BUC>
 A:Cross-references: GB:M64377
 C:Superfamily: olfactory receptor OR14
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 36.2%; Score 591; DB 2; Length 313;

Best Local Similarity 38.5%; Pred. No. 5.2e-43;

Matches 116; Conservative 56; Mismatches 129; Indels 0; Gaps 0;

QY 6 NOSTYDGEFLIGTSHQTDLVLFSAVNVYVALCGNVLLIFLIYDAGLHTPMYFELS 65
 Db 5 NOSTYDGEFLIGTSHQTDLVLFSAVNVYVALCGNVLLIFLIYDAGLHTPMYFELS 64
 QY 66 QLSIMDLMLVNCNIPKMAANFLSGRKSISFVCGIOIGFVSLVSGREGLLGLMAVDHYV 125
 Db 65 NLAEVDICYSATIPKMLQSFWEDECSISTYGCYIQLVYTFEATSDCYLLAAVMAVDHYV 124
 QY 126 AVSHPLHYPIILMNQVCLQITGSSMAFGIIDGVIQMWAAAMGLPYCGSRSDHFECEVOAL 185
 Db 125 AICPRLRYPIILMSQVCLLVALSYLMSINSVHTGFTFSLSCNSKNINHFECDDVPI 184
 QY 186 LKACADTSLFDTLIFACCVFMLLPFSIIMASTACILGAVLRIRSAQAMKALATCSSH 245
 Db 185 ISLSCSNNDIMIMLIVFVGFNLFTFVLIFFSYIYIMAILKMSSTAGRKFTSTCASH 244
 QY 246 LTAATLFFGAMFMYLPRRYRAPSHDKVASIFETVLTPLMPLIYSLRNREV 298
 Db 245 LTAATIRGTISTMILQPHSDNSEMKAVAFVGIYVIMPLNPLIYSLRNEV 297
 QY 306 L 306
 Db 305 L 305

RESULT 13

JC5200

chemoreceptor TB334 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-Aug-1999

C:Accession: JC5200; PC4302

R:Thomas, M.B.; Haines, S.L.; Akesson, R.A.

Gene 178, 1-5, 1996

A:Title: Chemoreceptors expressed in taste, olfactory and male reproductive tissues.

A:Reference number: JC5200; MUID:97080538; PMID:8921863

A:Accession: JC5200

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-311 <THO1>

A:Cross-references: GB:U50947; NID:g1256388; PIDN:AAC52909.1; PID:g1256389

A:Accession: PC4302

A:Status: preliminary

A:Molecule type: protein

A:Residues: 146-153;265-272 <THO2>

A:Experimental source: taste bud

C:Comment: This protein is coupled to a GTP-binding protein-mediated signal transduct

C:Genetics:

A:Gene: tb334

C:Superfamily: olfactory receptor OR14

C:Keywords: Olfaction; taste bud; transmembrane protein

F:24-47/Domain: transmembrane #status predicted <TM1>

F:56-77/Domain: transmembrane #status predicted <TM2>

F:99-118/Domain: transmembrane #status predicted <TM3>

F:138-162/Domain: transmembrane #status predicted <TM4>

F:195-217/Domain: transmembrane #status predicted <TM5>

F:236-258/Domain: transmembrane #status predicted <TM6>

F:271-291/Domain: transmembrane #status predicted <TM7>

```

Query Match          36.0%; Score 588; DB 2; Length 311;
Best Local Similarity 40.2%; Pred. No. 9.4e-43;
Matches 123; Conservative 53; Mismatches 130; Indels 0; Gaps 0;

OY 6 NOSYIDGFFLLGIFSHSOTDLVFSANVVFVALCGNVLLFLYLDAGLHTPMYFELS 65
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 NOSVSEFFLRNGISGFPEQOQLLYGFLCMYLVLTGCVLLIILALIGSPHLLTPMYFFLA 62

OY 66 OLSLMDLMVNCIVPKMANFLSGRKSISFVCGCIQIGFVSLVSEGLLGLIMAYDHYV 125
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 NLSFADMGILISTYTKMLFNVTOCHTISYICLQMTLPMFGLDSEFLVMAIDRYV 122

OY 126 AVSHPLHPIILMNQVRCIQITGSSWAFGIIDGVIOVAMAGLPYCGSRSDHFECEVAL 185
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 AICHPHLHSTIMNARICVLMILCMLFNVAVALFHTLLMARLSFCVGBIAHFECDVMSV 182

OY 186 LKLACADISLEPTLLFACCVFMLLPFSITMASYACILGAVLRISQAQMKALATCSGH 245
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 MRLSCSDTHVNLVLISGCGTVLWPFVSIYVHVFAYLRIGSSGSSKAFSTCSGH 242

OY 246 LTAVTLFGAAMFVTLRPRRYRAPSMDKVASIFETVLPMLPLIYSLRNGEVMGALRKG 305
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 LCVVCFVFGTLEFVSYLIFPSSVETTEKDVAAAAMTYVTPMLNPFIYSLRNDIKAKRL 302

OY 306 LDCRI 311
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 LSHRI 308

RESULT 14
S29707
Olfactory receptor OR5 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S29707; B37286
R:Ratting, K.; Krieger, J.; Strotmann, J.; Boekhoff, I.; Kubick, S.; Baumstark, C.; Breez
  Nature 361, 353-356, 1993
A:Title: Cloning and expression of odorant receptors.
A:Reference number: S29707; MUID:93149273; PMID:7678922
A:Accession: S29707
A:Molecule type: mRNA
A:Residues: 1-314 <RAT>
R:Buck, L.; Axel, R.
  Cell 65, 175-187, 1991
A:Title: A novel multigene family may encode odorant receptors: a molecular basis for od
  A:Reference number: A23701; MUID:91191556; PMID:1840504
A:Accession: B37286
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 193-236 <BUC>
A:Cross-references: GB:M64375; NID:g205811; PIDN:AAA41738.1; PID:g205812
C:Superfamily: olfactory receptor OR14
C:Keywords: G protein-coupled receptor; membrane protein

Query Match          35.7%; Score 583; DB 2; Length 314;
Best Local Similarity 38.7%; Pred. No. 2.5e-42;
Matches 115; Conservative 60; Mismatches 122; Indels 0; Gaps 0;

OY 6 NOSYIDGFFLLGIFSHSOTDLVFSANVVFVALCGNVLLFLYLDAGLHTPMYFELS 65
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 NOTVISOFLGLLPPREHOFYALFLSMYLTTLGNLITLILLLSHLTPMYFELS 64

OY 66 OLSLMDLMVNCIVPKMANFLSGRKSISFVCGCIQIGFVSLVSEGLLGLIMAYDHYV 125
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 65 NLSFEDLCFSSVTMPKLLQNNQSQVPSIPYAGCLSQIYFLFPGDLGNFLVAMAYDRYV 124

OY 126 AVSHPLHPIILMNQVRCIQITGSSWAFGIIDGVIOVAMAGLPYCGSRSDHFECEVAL 185
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 125 AICFPLHYSIMSPKLCVSLVLSWVLTTFHMLHTLLMARLSFCEVDVYIPHFCDMSAL 184

OY 186 LKLACADISLEPTLLFACCVFMLLPFSITMASYACILGAVLRISQAQMKALATCSGH 245
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 LKLACSDTRVNEVTLFVLSLFLVLPALITMSYVIVSILKVPSSQGIYKAFSSCGSH 244

```

```

OY 246 LTAVTLFGAAMFVTLRPRRYRAPSMDKVASIFETVLPMLPLIYSLRNGEVMGAL 302
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 LSVSLFETGTVLPYLCPSSNNSTYKETVMSLMTLVTPMLNPFIYSLRNDIKGAM 301

```

RESULT 15 JC5624

```

Olfactory receptor HPRCR92 - human
C:Species: Homo sapiens (man)
C:Date: 14-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 20-Jun-2000
C:Accession: JC5624; S58003
R:Vanderhaeghen, P.; Schumann, S.; Vassart, G.; Parmentier, M.
  Biochem. Biophys. Res. Commun. 237, 283-287, 1997
A:Title: Molecular cloning and chromosomal mapping of olfactory receptor genes expres
  A:Reference number: JC5624; MUID:97415789; PMID:9268701
A:Contents: Testis
A:Accession: JC5624
A:Molecule type: mRNA
A:Residues: 1-319 <VAN>
A:Cross-references: GB:Y10530; NID:g2792017; PIDN:CAA71558.1; PID:g2792018
  R:Vanderhaeghen, P.; Schumann, S.; Vassart, G.; Parmentier, M.
  submitted to the EMBL Data Library, July 1995
A:Description: Male germ cells from several mammalian species express a specific repe
  A:Reference number: S57995
A:Accession: S58003
A:Molecule type: mRNA
A:Residues: 126-282 <VAN>
A:Cross-references: EMBL:X89677; NID:g902337; PIDN:CAA61824.1; PID:g902338
C:Genetics:
A:Map position: 19p13.1
C:Function:
  A:Description: involved in control of sperm physiology
  C:Superfamily: olfactory receptor OR14
  C:Keywords: G protein-coupled receptor; transmembrane protein
  F:26-48/Domain: transmembrane #status predicted <TM1>
  F:58-79/Domain: transmembrane #status predicted <TM2>
  F:101-120/Domain: transmembrane #status predicted <TM3>
  F:140-164/Domain: transmembrane #status predicted <TM4>
  F:197-218/Domain: transmembrane #status predicted <TM5>
  F:237-260/Domain: transmembrane #status predicted <TM6>
  F:272-292/Domain: transmembrane #status predicted <TM7>

Query Match          35.4%; Score 578; DB 2; Length 319;
Best Local Similarity 38.0%; Pred. No. 6.8e-42;
Matches 113; Conservative 61; Mismatches 123; Indels 0; Gaps 0;

OY 6 NOSYIDGFFLLGIFSHSOTDLVFSANVVFVALCGNVLLFLYLDAGLHTPMYFELS 65
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 NDTQISEFLLLGFSQEPGLQPLFLFLSMYLVVLGMLLILATISDSHLTPMYFELS 64

OY 66 OLSLMDLMVNCIVPKMANFLSGRKSISFVCGCIQIGFVSLVSEGLLGLIMAYDHYV 125
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 65 NLSFADICVTSTTIPKMLNIOTONKVTYITACLMQMFILFAGFENFLSVAMAYDRYV 124

OY 126 AVSHPLHPIILMNQVRCIQITGSSWAFGIIDGVIOVAMAGLPYCGSRSDHFECEVAL 185
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 125 AICHPHVMVIMNPHLCGLVLASWTMSALYSLQILMVVRLSFCTALEIPHFCELMQV 184

OY 186 LKLACADISLEPTLLFACCVFMLLPFSITMASYACILGAVLRISQAQMKALATCSGH 245
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 IQLACSDSFLNMVITYFTVALGGGPIGLIYSYSKIISIIHAISSAGKTKAFSTCASH 244

OY 246 LTAVTLFGAAMFVTLRPRRYRAPSMDKVASIFETVLPMLPLIYSLRNGEVMGAL 302
    ||| ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 LSVSLFETGALIGVLLSSAATRNSSATASVMTVTPMLNPFIYSLRNDIKAL 301

```

Search completed: April 28, 2003, 14:02:43
Job time : 20 secs


```

OX NCBL_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BL/6; TISSUE=OLFACTORY EPITHELIUM;
RX MEDLINE=99091050; Pubmed=9875846;
RA Krautwurst D., Yau K.W., Reed R.R.;
RT "Identification of ligands for olfactory receptors by functional
   expression of a receptor library.";
RL Cell 95:917-926(1998).
DR EMBL: AF102535; AAD13327.1; -
MGD: MGI:1333785; Olf56.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECP_FL_2; 1.
KW Receptor.
FT NON_TER
FT SEQUENCE
SQ
  1 223
  223 223
  24591 MW; 8D421B5850475F93 CRC64;

Query Match
Best Local Similarity 63.6%; Score 1040; DB 11; Length 223;
Matches 199; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 65 SLSLMDMLVNCNIVPKMANFLSGRKSISFVCGIQTGFVSLVSGELLGLMAYDHY 124
DB 1 SLSLMDMLVNCNIVPKMANFLSGRKSISFVCGIQTGFVSLVSGELLGLMAYDRY 60

QY 125 VAVSHPLHYPLIMNORVCIQITSSNAFGIIDVYIOMVAMGLPYGSRVDHFFCEVQA 184
DB 61 VAVSHPLHYPLIMNORVCIQITSSNAFGIIDVYIOMVAMGLPYGSRVDHFFCEVPA 120

QY 185 LKLAACADTSLFPTLLFACCVFMLLPSTIMASYACIIIGAVLRISAQAKKALATCCS 244
DB 121 LKLAACADTSLFPTLLFACCVFMLLPSTITVSYARILIGTVLRMSAQSRRKALATCCS 180

QY 245 HLTAVTLEFGAMFMYLRPRRYRAPSHDKVASIFYTVLPMLN 287
DB 181 HLTAVTLEFGAMFMYLRPRRYRAPSHDKVASIFYTVLPMLN 223

RESULT 5
Q8VGD7
ID 08VGD7 PRELIMINARY; PRT; 317 AA.
AC 08VGD7;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-JUN-2002 (Tremblrel. 20, Last sequence update)
DE Olfactory receptor MOR277-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBL_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang X., Firestein S.J.;
RT "The olfactory receptor gene superfamily of the mouse.";
RL Nat. Neurosci. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Adams M.;
RN Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AY073213; AAL60876.1; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS50262; G_PROTEIN_RECP_FL_2; 1.
KW Receptor.
FT NON_TER
FT SEQUENCE
SQ
  317 AA; 35453 MW; 331CD9B8C0CADAB1 CRC64;

Query Match
Best Local Similarity 61.8%; Score 1010.5; DB 11; Length 317;
Matches 199; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

```

```

Matches 192; Conservative 46; Mismatches 71; Indels 1; Gaps 1;

QY 4 WV-NOSYDGFLLGIFSHSOTDLVFSAVMVEFVAGCNVLLIFLIYDAGATPMYF 62
DB 3 WAGNOLIFSHFVLGLFTSPPLHFLFSTIMVFWALVSLNGMILLILMDSLRHPMF 62

QY 63 FLSLSLMDMLVNCNIVPKMANFLSGRKSISFVCGIQTGFVSLVSGELLGLMAYD 122
DB 63 FLSLSLMDMLVNCNIVPKMANFLSGRKSISFVCGIQTGFVSLVSGELLGLMAYD 122

QY 123 HYAVSHPLHYPLIMNORVCIQITSSNAFGIIDVYIOMVAMGLPYGSRVDHFFCEY 182
DB 123 HYAVSHPLHYPLIMNORVCIQITSSNAFGIIDVYIOMVAMGLPYGSRVDHFFCEY 182

QY 183 QALLKACADTSLFPTLLFACCVFMLLPSTIMASYACIIIGAVLRISAQAKKALATC 242
DB 183 PAVKLACADTSLFPTLLFACCVFMLLPSTITVSYARILIGTVLRMSAQSRRKALATC 242

QY 243 SSHLTAVTLEFGAMFMYLRPRRYRAPSHDKVASIFYTVLPMLNPLIYSLRGEVYAGAL 302
DB 243 SSHLTAVTLEFGAMFMYLRPRRYRAPSHDKVASIFYTVLPMLNPLIYSLRGEVYAGAL 302

QY 303 RKGLDRCRG 312
DB 303 RKLLGRCPG 312

RESULT 6
Q96R30
ID Q96R30 PRELIMINARY; PRT; 216 AA.
AC Q96R30;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Olfactory receptor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fuchs T., Malecova B., Linhart C., Sharan R., Khen M., Herwig R.,
RA Shmulevich D., Elkon R., Steinfath M., O'Brien J.K., Radloff U.,
RA Lehrach H., Olander 2., Glusman G., Lancet D., Shamir R.;
RT "DEPOG: A Practical Scheme for Deciphering Families of Genes.";
RT Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF399614; AAK95099.1; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECP_FL_2; 1.
KW Receptor.
FT NON_TER
FT NON_TER
FT SEQUENCE
SQ
  216 AA; 24042 MW; B7F41AF784E61D20 CRC64;

Query Match
Best Local Similarity 59.7%; Score 975; DB 4; Length 216;
Matches 185; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

QY 69 LMDLMLVNCNIVPKMANFLSGRKSISFVCGIQTGFVSLVSGELLGLMAYDHYAVS 128
DB 1 LMDLMLVNCNIVPKMANFLSGRKSISFVCGIQTGFVSLVSGELLGLMAYDRYAVS 60

QY 129 HPLHYPLIMNORVCIQITSSNAFGIIDVYIOMVAMGLPYGSRVDHFFCEVQALKL 188
DB 61 HPLHYPLIMNORVCIQITSSNAFGIIDVYIOMVAMGLPYGSRVDHFFCEVQALKL 120

QY 189 ACADTSLFPTLLFACCVFMLLPSTIMASYACIIIGAVLRISAQAKKALATCCSHLTA 248
DB 121 ACYDTSLEFKVIFACCVFMLLPSTITVSYARILIGTVLRMSAQSRRKALATCCSHLTA 180

QY 249 VTLFYGAMFMYLRPRRYRAPSHDKVASIFYTVLP 284

```

Db 181 VTLEFYGAAMFYILRPRHYRAPSHDKVASIFYVLTLP 216

RESULT 7

Q8VF87 PRELIMINARY; PRT; 315 AA.

AC 08VF87;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Olfactory receptor MOR279-2.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RA SEQUENCE FROM N.A.
RA Zhang X., Firestein S.J.;
RT "The olfactory receptor gene superfamily of the mouse."
RL Nat. Neurosci. 0:0-0(2002).

RP SEQUENCE FROM N.A.

RA Adams M.;
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
RL EMBL; AY073647; AAL61310.1;
DR InterPro: IPR000276; GPCR_Rhodopsn.

DR Pfam: PF00001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRHOOPS.

DR PROSITE; PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.

DR PROSITE; PS0262; G_PROTEIN_RECPEP_FL_2; 1.

KW Receptor.

SQ SEQUENCE 315 AA; 35208 MW; 05FE088ACD114099 CRC64;

Query Match 56.4%; Score 921; DB 11; Length 315;
Best Local Similarity 55.5%; Pred. No. 2.8e-75;

Matches 172; Conservative 53; Mismatches 83; Indels 2; Gaps 1;

QY 1 MGRWVNSYTDGFFLLGIFSHSQTDVLFSAVWVFTVALCGNVLLIFLYLDAGLHTPM 60

Db 3 MDMNNWTRNSDFILGLFDHSPHFFFLIGIFFMATIGNSINVIILYLDAGLHTPM 62

QY 61 YFPLSLSLMDLVCNIVPKMANFLSGKRSIFVCGCIQIGFVSVLVSSEGLLGMA 120

Db 63 YILLSLSLMDLICTVQAFNPLSGNKSISMVCGCIQIFVSVLGAECFLAAMA 122

QY 121 YDHVAVSHPLHYPIILMNRVCQITGSSWAFGIIDGVIQMVAMGLPYCGSSVDFHC 180

Db 123 YDRYVAICPLRPIILMSDKICGLMAASSWVLSLDGILEVAALSFYCGAREIPHFPC 182

QY 181 EVOALKLACADTSLPDTLLFACCVFMLLPFSIIIMASYACILGAVLRISAQAMKALA 240

Db 183 DVPALLTSCNSLIFERITFCVCVIMLTLPVAILIATYRVLTYLHMSASRSRKAPA 242

QY 241 TCGSHLTAVTLFYGAMFMYLRPRRYRAPSHDKVASIFYVLTLPMLNPLIYSRNEVWG 300

Db 243 TCSHLMVVMYGYGAMFYMRPSSGSRPTQDKIVSAFYITLPLNPLIYSRNEVAR 302

QY 301 ALRK--GLDR 308

Db 303 AFMKVLGIDK 312

RESULT 8

Q8VFA2 PRELIMINARY; PRT; 313 AA.

AC 08VFA2;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Olfactory receptor MOR279-1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RA SEQUENCE FROM N.A.
RA Zhang X., Firestein S.J.;
RT "The olfactory receptor gene superfamily of the mouse."
RL Nat. Neurosci. 0:0-0(2002).

RP SEQUENCE FROM N.A.

RA Adams M.;
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
RL EMBL; AY073632; AAL61295.1;
DR InterPro: IPR000276; GPCR_Rhodopsn.

DR Pfam: PF00001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRHOOPS.

DR PROSITE; PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.

DR PROSITE; PS0262; G_PROTEIN_RECPEP_FL_2; 1.

KW Receptor.

SQ SEQUENCE 313 AA; 35116 MW; 623926357774A7C0 CRC64;

Query Match 54.7%; Score 893; DB 11; Length 313;
Best Local Similarity 52.6%; Pred. No. 9.3e-73;

Matches 161; Conservative 57; Mismatches 88; Indels 0; Gaps 0;

QY 1 MGRWVNSYTDGFFLLGIFSHSQTDVLFSAVWVFTVALCGNVLLIFLYLDAGLHTPM 60

Db 1 MDMNNWTRNSDFILGLFDHSPHFFFLIGIFFMATIGNSINVIILYLDAGLHTPM 60

QY 61 YFPLSLSLMDLVCNIVPKMANFLSGKRSIFVCGCIQIGFVSVLVSSEGLLGMA 120

Db 61 YILLSLSLMDLICTVQAFNPLSGNKSISMVCGCIQIFVSVLGAECFLAAMA 120

QY 121 YDHVAVSHPLHYPIILMNRVCQITGSSWAFGIIDGVIQMVAMGLPYCGSSVDFHC 180

Db 121 YDRYVAICPLRPIILMSDKICGLMAASSWVLSLDGILEVAALSFYCGAREIPHFPC 180

QY 181 EVOALKLACADTSLPDTLLFACCVFMLLPFSIIIMASYACILGAVLRISAQAMKALA 240

Db 181 DIPALLTSCNSLIFERITFCVCVIMLTLPVAILIATYRVLTYLHMSASRSRKAPA 240

QY 241 TCGSHLTAVTLFYGAMFMYLRPRRYRAPSHDKVASIFYVLTLPMLNPLIYSRNEVWG 300

Db 241 TCSHLMVVMYGYGAMFYMRPSSGSRPTQDKIVSAFYITLPLNPLIYSRNEVAR 300

QY 301 ALRKGL 306

Db 301 AFMKVL 306

RESULT 9

Q8VGT5 PRELIMINARY; PRT; 317 AA.

AC 08VGT5;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Olfactory receptor MOR274-1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RA SEQUENCE FROM N.A.

RA Zhang X., Firestein S.J.;

RT "The olfactory receptor gene superfamily of the mouse."
RL Nat. Neurosci. 0:0-0(2002).

RP SEQUENCE FROM N.A.

RA Adams M.;

RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.

DR EMBL; AY073061; AAL60724.1;

DR InterPro: IPR000276; GPCR_Rhodopsn.

DR Pfam: PF00001; 7tm_1; 1.

Db 245 SHLAVVLYGGAISFYSMORASARPLGDRATSIFFYIVTPEFNPVLSLRNREVTALK 304
 QY 304 KGLDR 308
 Db 305 KMLER 309

RESULT 12 Q8VGX2 PRELIMINARY: PRT: 312 AA.

AC Q8VGX2: 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Olfactory receptor MOR270-1.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang X., Firestein S.J.;
 RT "The olfactory receptor gene superfamily of the mouse."
 RL Nat. Neurosci. 0:0-0(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Adams M.;
 RT Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY073019; AAL60682.1; -;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECCEP_F1_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECCEP_F1_2; 1.
 DR KX Receptor.
 SQ SEQUENCE 312 AA: 35391 MW: 62504A0407AFC976 CRC64;

Query Match 52.5%; Score 858; DB 11; Length 312;
 Best Local Similarity 54.2%; Pred. No. 1.3e-69;
 Matches 166; Conservative 55; Mismatches 83; Indels 2; Gaps 2;

QY 1 MCRWVNSTDGFLLIGFSHSOTDLVFSAVWVFTVALCGVLLIFLITLDAGLHTPM 60
 Db 1 MEKW-NQSSD-FLLGLLPQNOTGLMLLIIVFSLCNGSGVILHILRVDPLHTPM 58
 QY 61 YFFLSQSLMDLMLVCNIPKMAANFLSGKRSISFVGCIOIGFFVSLVSGGLLGLMA 120
 Db 59 YFFLSQSLMDLMLYSTVTPKMAFNFLSGKRSISFVGCIOIGFFVSLVSGGLLGLMA 118
 QY 121 YDHVAVSHPLHYPIIMNORVCLQTGSSWAFGIIDGVIQWYAANGLPYCGSRSDHFC 180
 Db 119 YDFVAICHPLHYPIIRMSKIMCLMIGSWILGINSLSAHYVAALHIPYCHRSINHFPC 178
 QY 181 EVOALILKLACADTSLFDTLFACVFMLLPFSIIMASYACILGAVLRIRSAQWAKKALA 240
 Db 179 DVPAMPLPLACMNTWYEVFVSTSLFLLPLGLITASGRVLFVAFHRSKEGKKKAT 238
 QY 241 TCSSHLTAVTLEFGAEMYLRRPRYRAPSHDKVASIEFTVLTPLMLNPLIYSLRGEVWG 300
 Db 239 TCSHLTVVTFYAFVYVYLRPRSLRSPTEDEKILAVFTIILPMLNPLIYSLRKEVVG 298
 QY 301 ALRKL 306
 Db 299 AMTRVL 304

RESULT 13
 Q8VF05 PRELIMINARY: PRT: 312 AA.
 AC Q8VF05: 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Olfactory receptor MOR271-1.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang X., Firestein S.J.;
 RT "The olfactory receptor gene superfamily of the mouse."
 RL Nat. Neurosci. 0:0-0(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Adams M.;
 RT Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY073736; AAL61399.1; -;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_RHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECCEP_F1_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECCEP_F1_2; 1.
 DR KX Receptor.
 SQ SEQUENCE 312 AA: 35536 MW: 2F11DC131EDE829D CRC64;

Query Match 52.4%; Score 856; DB 11; Length 312;
 Best Local Similarity 53.6%; Pred. No. 2e-69;
 Matches 164; Conservative 57; Mismatches 83; Indels 2; Gaps 2;

QY 1 MCRWVNSTDGFLLIGFSHSOTDLVFSAVWVFTVALCGVLLIFLITLDAGLHTPM 60
 Db 1 MEKW-NQSSD-FLLGLLPQNOTGLMLLIIVFSLCNGSGVILHILRVDPLHTPM 58
 QY 61 YFFLSQSLMDLMLVCNIPKMAANFLSGKRSISFVGCIOIGFFVSLVSGGLLGLMA 120
 Db 59 YFFLSQSLMDLMLYSTVTPKMAFNFLSGKRSISFVGCIOIGFFVSLVSGGLLGLMA 118
 QY 121 YDHVAVSHPLHYPIIMNORVCLQTGSSWAFGIIDGVIQWYAANGLPYCGSRSDHFC 180
 Db 119 YDFVAICHPLHYPIIRMSKIMCLMIGSWILGINSLSAHYVAALHIPYCHRSINHFPC 178
 QY 181 EVOALILKLACADTSLFDTLFACVFMLLPFSIIMASYACILGAVLRIRSAQWAKKALA 240
 Db 179 DVPAMPLPLACMNTWYEVFVSTSLFLLPLGLITASGRVLFVAFHRSKEGKKKAT 238
 QY 241 TCSSHLTAVTLEFGAEMYLRRPRYRAPSHDKVASIEFTVLTPLMLNPLIYSLRGEVWG 300
 Db 239 TCSHLTVVTFYAFVYVYLRPRSLRSPTEDEKILAVFTIILPMLNPLIYSLRKEVVG 298
 QY 301 ALRKL 306
 Db 299 AMTRVL 304

RESULT 14 Q8VGD8 PRELIMINARY: PRT: 315 AA.

AC Q8VGD8: 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Olfactory receptor MOR281-1.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang X., Firestein S.J.;
 RT "The olfactory receptor gene superfamily of the mouse."
 RL Nat. Neurosci. 0:0-0(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Adams M.;
 RT Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY073212; AAL60875.1; -;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:59:46 ; Search time 14 Seconds
(without alignments)
662.016 Million cell updates/sec

Title: US-09-634-109d-2

Perfect score: 1634

Sequence: 1 MGRWVNSQSYDGFLLGIFS.....GEVWGLRKGRDRCRIGSOH 315

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

262574

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*

5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep:*

6: /cgn2_6/ptodata/1/iaa/Deckfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 869 | 53.2 | 316 | 2 | US-08-827-291A-2 |
| 2 | 636.5 | 39.0 | 321 | 4 | US-08-748-506-20 |
| 3 | 635.5 | 38.9 | 321 | 4 | US-08-748-506-18 |
| 4 | 633.5 | 38.8 | 321 | 4 | US-08-748-506-12 |
| 5 | 632.5 | 38.7 | 321 | 4 | US-08-748-506-10 |
| 6 | 630.5 | 38.6 | 321 | 4 | US-08-748-506-11 |
| 7 | 630.5 | 38.6 | 321 | 4 | US-08-748-506-13 |
| 8 | 630.5 | 38.6 | 321 | 4 | US-08-748-506-19 |
| 9 | 610.5 | 37.4 | 296 | 2 | US-08-467-948A-2 |
| 10 | 610.5 | 37.4 | 296 | 2 | US-08-467-948A-2 |
| 11 | 588 | 36.0 | 309 | 3 | US-08-988-876-5 |
| 12 | 571 | 34.9 | 333 | 3 | US-08-988-876-6 |
| 13 | 570 | 34.9 | 314 | 3 | US-08-988-876-7 |
| 14 | 534 | 32.7 | 284 | 1 | US-08-118-270-61 |
| 15 | 534 | 32.7 | 284 | 5 | PCT-US93-08528-61 |
| 16 | 533 | 32.6 | 222 | 2 | US-08-467-948A-27 |
| 17 | 533 | 32.6 | 222 | 3 | US-08-467-948A-27 |
| 18 | 529.5 | 32.4 | 286 | 1 | US-08-118-270-65 |
| 19 | 529.5 | 32.4 | 286 | 5 | PCT-US93-08528-65 |
| 20 | 501 | 30.7 | 327 | 4 | US-08-748-506-22 |
| 21 | 501 | 30.7 | 327 | 4 | US-08-748-506-23 |
| 22 | 499.5 | 30.6 | 284 | 1 | US-08-118-270-67 |
| 23 | 499.5 | 30.6 | 284 | 5 | PCT-US93-08528-67 |
| 24 | 497.5 | 30.4 | 293 | 1 | US-08-118-270-60 |
| 25 | 497.5 | 30.4 | 293 | 5 | PCT-US93-08528-60 |
| 26 | 480 | 29.4 | 327 | 4 | US-08-748-506-14 |
| 27 | 479 | 29.3 | 327 | 4 | US-08-748-506-24 |

| | | | | | | |
|----|-------|------|-----|---|-------------------|-------------------|
| 28 | 468.5 | 28.7 | 274 | 1 | US-08-118-270-69 | Sequence 69, Appl |
| 29 | 468.5 | 28.7 | 274 | 5 | PCT-US93-08528-69 | Sequence 69, Appl |
| 30 | 458.5 | 28.1 | 273 | 1 | US-08-118-270-63 | Sequence 63, Appl |
| 31 | 458.5 | 28.1 | 273 | 5 | PCT-US93-08528-63 | Sequence 63, Appl |
| 32 | 457 | 28.0 | 247 | 1 | US-08-465-980-3 | Sequence 3, Appl1 |
| 33 | 457 | 28.0 | 247 | 2 | US-09-053-303-3 | Sequence 3, Appl1 |
| 34 | 457 | 28.0 | 247 | 5 | PCT-US95-07093-3 | Sequence 3, Appl1 |
| 35 | 457 | 28.0 | 247 | 5 | PCT-US95-07093-3 | Sequence 3, Appl1 |
| 36 | 447.5 | 27.4 | 277 | 1 | US-08-118-270-68 | Sequence 68, Appl |
| 37 | 447.5 | 27.4 | 277 | 5 | PCT-US93-08528-68 | Sequence 68, Appl |
| 38 | 445 | 27.2 | 277 | 1 | US-08-118-270-62 | Sequence 62, Appl |
| 39 | 445 | 27.2 | 277 | 5 | PCT-US93-08528-62 | Sequence 62, Appl |
| 40 | 443 | 27.1 | 269 | 1 | US-08-118-270-64 | Sequence 64, Appl |
| 41 | 443 | 27.1 | 269 | 5 | PCT-US93-08528-64 | Sequence 64, Appl |
| 42 | 443 | 26.5 | 275 | 1 | US-08-118-270-66 | Sequence 66, Appl |
| 43 | 433 | 26.5 | 275 | 5 | PCT-US93-08528-66 | Sequence 66, Appl |
| 44 | 347 | 21.2 | 320 | 1 | US-08-465-980-2 | Sequence 2, Appl1 |
| 45 | 347 | 21.2 | 320 | 2 | US-09-053-303-2 | Sequence 2, Appl1 |

ALIGNMENTS

RESULT 1
US-08-827-291A-2
Sequence 2, Application US/08827291A
Patent No. 5874243
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
TITLE OF INVENTION: NOVEL OLRCC15 RECEPTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY:
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/827,291A
FILING DATE: 28-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: King, William T
REGISTRATION NUMBER: 30,954
REFERENCE/DOCKET NUMBER: GP50001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5015
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-827-291A-2
Query Match 53.2%; Score 869; DB 2; Length 316;
Best Local Similarity 54.2%; Pred. No. 2.1e-72;
Matches 162; Conservative 47; Mismatches 90; Indels 0; Gaps 0;
QY 6 NSQYDGFLLGIFSHSQNDLVLSAVMVETVALCGNVLIFLIYLDAGLHRTMYFFLS 65

```

Db      5 NSTFNSDFELGJFNHSPHTFFELVLAIFSVAFMGNVAVLLITYDTQLHTPMYLLS 64
QY      66 QLSMDMLVCNIVPKKAAANFLSGRKSISFVCGGICGIFVSLVSGEGLLGLMAYHY 125
        65 QLSMDMLVCNIVPKKAAANFLSGRKSISMACGCTQIFFTSLGSGCPLLAWAYRYRT 124
QY      126 AVSHPLHYPLIMNORVCIQTGSSMAGIIDGVLIOMVAANGLPYCGSRVDHFECEVAL 185
        125 AICHPLEVTNLSMSPKICGLMTAFSMILGSTDGIIYAVATFSEYSGSRREIAHFCELPSTL 184
Db      186 LKACADTSLFDTLFACCVFMLLPESIIMASVACILGAVLRIRSAQANKKALATGSH 245
        185 LILSCNDTSIFEKVIFCSITVMTLVPVAILIISYAGVILAVIHMGSEGRKRAFTCSSH 244
QY      246 LTAVTLEFYGAAMEYLRPRRYRAPSIRKVASIETVTLTPMLNPLIYSLRNGEVAALRK 304
        245 LMYGMFYGAGLFPMTIOPTSDRSPOTDKIVSVFTTILTPMLNPLIYSLRNGEVAALRK 303

```

RESULT 2

```

US-08-748-506-20
; Sequence 20, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Romnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Volt & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5700
; TELEFAX: 312-616-5600
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-748-506-20

```

```

Query Match      39.0%; Score 636.5; DB 4; Length 321;
Best Local Similarity 42.8%; Pred. No. 5.1e-51;
Matches 130; Conservative 54; Mismatches 119; Indels 1; Gaps 1;

```

```

QY      4 WVNOSYTDGFFLLGIFHSQTD-LVLFSAVMVVFVALCGNVLLIFLYLDAGLHTPMYF 62
        8 WOENSLYKRFAPFAKSEYGECEFLFTLLMLFLVSLTGNLILALICTSPLTPMYF 67
Db      63 FLSQSLMDMLVCNIVPKKAAANFLSGRKSISFVCGGICGIFVSLVSGEGLLGLMAYD 122
        68 FLANSLLEIGTCTSVIRKMQSLVSEARGISMEGACSMFFIFFGIRECCLLAAMAFD 127

```

```

QY      123 HYVAVSHPLHYPLIMNORVCIQTGSSMAGIIDGVLIOMVAANGLPYCGSRVDHFECEV 182
        128 RYMAICSPLEHYATRMRSKGVCAVLAIVSWNGCIVGLQNTFISLNFQCGCEIDHFFCDL 187
QY      183 QALIKLACADTSLFDTLFACCVFMLLPESIIMASVACILGAVLRIRSAQANKKALATC 242
        188 PLLALACGDTQSNENAEIFVAVNLICFSPPLLISSVRLIVAVLVMPSPGGRKALSTC 247
QY      243 SSHLTAVTLEFYGAAMEYLRPRRYRAPSIRKVASIETVTLTPMLNPLIYSLRNGEVAAL 302
        248 SSHLTAVTLEFYGSATVYLRKSSHSPGVKDLALFYTSVSMNLPIIYSLRNGEVAAL 307
QY      303 RKGL 306
        308 RRTL 311

```

RESULT 3

```

US-08-748-506-18
; Sequence 18, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Romnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Volt & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5700
; TELEFAX: 312-616-5600
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-748-506-18

```

```

Query Match      38.9%; Score 635.5; DB 4; Length 321;
Best Local Similarity 43.1%; Pred. No. 6.3e-51;
Matches 131; Conservative 50; Mismatches 122; Indels 1; Gaps 1;

```

```

QY      4 WVNOSYTDGFFLLGIFHSQTD-LVLFSAVMVVFVALCGNVLLIFLYLDAGLHTPMYF 62
        8 WOENSLYKRFAPFAKSEYGECEFLFTLLMLFLVSLTGNLILALICTSPLTPMYF 67
Db      63 FLSQSLMDMLVCNIVPKKAAANFLSGRKSISFVCGGICGIFVSLVSGEGLLGLMAYD 122
        68 FLANSLLEIGTCTSVIRKMQSLVSEARGISMEGACSMFFIFFGIRECCLLAAMAFD 127
QY      123 HYVAVSHPLHYPLIMNORVCIQTGSSMAGIIDGVLIOMVAANGLPYCGSRVDHFECEV 182
        128 RCMATCSPLHYATRMRSREVCAHLAIVSMGMCIVSLGQTNFISLFCGPCIIDHFFCDL 187

```


QY 303 RK 304
Db 308 RR 309

RESULT 8
US-08-748-506-19

Sequence 19, Application US/08748506
Patent No. 6159707

GENERAL INFORMATION:

APPLICANT: Ronnett et al.

TITLE OF INVENTION: NOVEL SPERM RECEPTORS

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Leydig, Voigt & Mayer, Ltd.

STREET: Two Prudential Plaza, Suite 4900

CITY: Chicago

STATE: IL

COUNTRY: US

ZIP: 60601-6780

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/748,506

FILING DATE: 08-NOV-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/033,751

FILING DATE: 09-NOV-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

REFERENCE/DOCKET NUMBER: 74940

TELEPHONE: 312-616-5600

TELEFAX: 312-616-5700

INFORMATION FOR SEQ. ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 321 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-748-506-19

Query Match 38.6%; Score 630.5; DB 4; Length 321;

Best Local Similarity 41.8%; Pred. No. 1.8e-50;

Matches 130; Conservative 53; Mismatches 127; Indels 1; Gaps 1;

QY 4 WVNQSYDGFLLGIFSHSQTDLVLFSAVMVETVALGCVNLLIFLIYDAGLHPMYF 62

Db 8 WQENSLSVKKFAAKTSFVGECEFLTLILMLFVSLTGNALIALAACPSPSLHPMYF 67

QY 63 FLSQSLMDLMLVNCNPKMANFLSGKRSISFVGGCIQIGFVSLVSGELLGLMAYD 122

Db 68 FLANLSLEIGYCSVPKMLQSLVSEARISREGCATOMFFFTFEGITCCLLAANA 127

QY 123 HYAAVSHPLHYPLIMNORVCLQITGSSMAFGIIDVYQVMAANGLPYCGRSVDHFCFV 162

Db 128 RCMGICSPLEHYARMSREVCANHLAIVSMGMCIVGLQITNNIISLNCGCEIDHFCFV 187

QY 183 QALIKLACADTSLEFDLILFACVFMILLPFSIMASYACILGAVLRIRSOAKKALATC 242

Db 188 PLLALACGTSQNEAALIFPAALICISPLVILYSVRLVAVLWVPSGEGRHKALSTC 247

QY 243 SSHLTAVTLFYGAMFMYLPRRYRAPSHDKVASIFVTYVLPMLNPLIYSLRNGEYV 302

Db 248 SSHLLVTLFYGVSFTYLPKPSHSGMDKLLALFYTAVTSMNLPIIYSLRNGEYV 307

QY 303 RKGLDRCRIGS 313

Db 308 RRTLDLKRIMS 318

RESULT 9
US-08-467-948A-2

Sequence 2, Application US/08467948A
Patent No. 5998164

GENERAL INFORMATION:

APPLICANT: LI, YI

APPLICANT: CAO, JIAN

APPLICANT: NI, JIAN

APPLICANT: GENTZ, REINER

APPLICANT: BULT, CAROL J.

APPLICANT: SUTTON III, GRANGER G.

APPLICANT: ROSEN, CRAIG A.

TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein

TITLE OF INVENTION: Coupled Receptor GPR2

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 NEW YORK AVE., NW, SUITE 600

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,948A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/04079

FILING DATE: 30-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: STEFFE, ERIC K.

REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM

TELEPHONE: 202-371-2540

TELEFAX: 202-371-2600

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 296 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-467-948A-2

Query Match 37.4%; Score 610.5; DB 2; Length 296;

Best Local Similarity 41.9%; Pred. No. 1.1e-48;

Matches 119; Conservative 56; Mismatches 108; Indels 1; Gaps 1;

QY 6 NQSYTDFEFLGIFSHSQTDLVLFSAVMVETVALGCVNLLIFLIYDAGLHPMYF 65

Db 8 NQWTFEFLGIFSHSQTDLVLFSAVMVETVALGCVNLLIFLIYDAGLHPMYF 67

QY 66 QLSMDLMLVNCNPKMANFLSGKRSISFVGGCIQIGFVSLVSGELLGLMAYDHY 125

Db 68 HLAVNIAVACNTVPQMLVNLHPAKIPSPAGCMTDLFLPSAHPFBCLLILMSYDRYV 127

QY 126 AVSHPLHYPLIMNORVCLQITGSSMAFGIIDVYQVMAANGLPYCGRSVDHFCFV 165

Db 128 AICHPLHYPLIMNORVCLQITGSSMAFGIIDVYQVMAANGLPYCGRSVDHFCFV 187

QY 186 LKLAACADTSLEFDLILFACVFMILLPFSIMASYACILGAVLRIRSOAKKALATC 245

Db 188 LRLACADTWLNQVIFEACGFIIVGFLCLVLSYSHLGLILQISGEGRAKAFSTCSSH 247

QY 246 LTAVTLFYGAMFMYLPRRYRAPSHDKVASIFVTYVLPMLNPLIYSLRNGEYV 302

Db 308 RRTLDLKRIMS 318

Db 248 LCYVGLFFGSAIWMYAPKSRHPEEQKVLFLLIQFLSTPMKLP 291

RESULT 10

US-08-467-947A-2
Sequence 2, Application US/08467947A

Patent No. 6090575

GENERAL INFORMATION:

APPLICANT: LI, YI

APPLICANT: CAO, LIANG

APPLICANT: NI, JIAN

APPLICANT: GEMTZ, REINER

APPLICANT: BUIE, CAROL J.

APPLICANT: SUTTON III, GRANGER G.

APPLICANT: ROSEN, CRAIG A.

TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein

TITLE OF INVENTION: Coupled Receptor GPR1

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P. L. L. C.

STREET: 1100 NEW YORK AVE., NW, SUITE 600

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467, 947A

CLASSIFICATION: 435

FILING DATE: 06-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/04079

FILING DATE: 30-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: STEFFE, ERIC K.

REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 296 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-467-947A-2

Query Match 37.4%; Score 610.5; DB 3; Length 296;

Best Local Similarity 41.9%; Pred. No. 1.1e-48;

Matches 113; Conservative 56; Mismatches 108; Indels 1; Gaps 1;

6 NOSTYDEFFLIGIFSHSQTDLVLFSAVWVFTVALCGNVLLIFLIYDAGLHTPMYFELS 65

8 NQWTEBELLGFLGFLGPRIGMLRGLFSLFVFTLLGNGTLLGLISDSRLHTPMYFELS 67

66 QLSLMDLMLVNIYPKMAANFLSGRKSISFVGCIGIQGFVSLVSGGLLGLMAVHYV 125

68 HLAIVNIYACNTPVQMLVNLHPAKPIFAGCMTLDLFLFSAHTTECLLVLMSSYDRYV 127

126 AVSPRLPHPIIMNORVCLQITGSSMAGIIGDGVQIOMVAMMGLPYCGSRSDHFCCEVAL 185

128 AICPLRLFIIMYKVCITLIGITSTGCSLLAMVHSLILRPLPGPREINHFCEILISV 187

186 LKLAACDLSFDTLFACVFMLLPFSSIIYASYACILGAVLRIRSAQAMKALATCSSH 245

188 LRLACADTWMNOVYIFECMFIIVGPLCLVLYVSHILGILRIQSGCRKRASTCSSH 247

246 LTAIVLEFYGAAMPYLRPRRYRAPSDDKVASIFTYVL-TPMLNP 288

Db 248 LCYVGLFFGSAIWMYAPKSRHPEEQKVLFLLIQFLSTPMKLP 291

RESULT 11

US-08-988-876-5
Sequence 5, Application US/08988876

Patent No. 6063596

GENERAL INFORMATION:

APPLICANT: Lai, Preeti

APPLICANT: Bandman, Olga

APPLICANT: Hillman, Jennifer L.

APPLICANT: Yue, Henry

TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED

TITLE OF INVENTION: WITH IMMUNE RESPONSE

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/988, 876

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PE-0441 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 309 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: Genbank

CLONE: 1314667

US-08-988-876-5

Query Match 36.0%; Score 588; DB 3; Length 309;

Best Local Similarity 38.0%; Pred. No. 1.4e-16;

Matches 113; Conservative 64; Mismatches 120; Indels 0; Gaps 0;

6 NOSTYDEFFLIGIFSHSQTDLVLFSAVWVFTVALCGNVLLIFLIYDAGLHTPMYFELS 65

5 NQWTEBELLGFLGFLGPRIGMLRGLFSLFVFTLLGNGTLLGLISDSRLHTPMYFELA 64

66 QLSLMDLMLVNIYPKMAANFLSGRKSISFVGCIGIQGFVSLVSGGLLGLMAVHYV 125

68 HLAIVNIYACNTPVQMLVNLHPAKPIFAGCMTLDLFLFSAHTTECLLVLMSSYDRYV 127

126 AVSPRLPHPIIMNORVCLQITGSSMAGIIGDGVQIOMVAMMGLPYCGSRSDHFCCEVAL 185

128 AICPLRLFIIMYKVCITLIGITSTGCSLLAMVHSLILRPLPGPREINHFCEILISV 187

186 LKLAACDLSFDTLFACVFMLLPFSSIIYASYACILGAVLRIRSAQAMKALATCSSH 245

188 LRLACADTWMNOVYIFECMFIIVGPLCLVLYVSHILGILRIQSGCRKRASTCSSH 247

246 LTAIVLEFYGAAMPYLRPRRYRAPSDDKVASIFTYVL-TPMLNP 288

Fri, May .9 14:21:33 2003

us-09-634-109d-2.ra1

Page 9

[illegible]

Search completed: April 28, 2003, 14:03:04
Job time : 15 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 20:27:35 ; Search time 2661 Seconds

(without alignments)
10368.080 Million cell updates/sec

Title: US-09-634-109D-1

Perfect score: 948

Sequence: 1 atggagagatggtgaccca.....ggattgcagccagcactga 948

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hg_hum.*
31: em_hg_inv.*
32: em_hg_other.*
33: em_hg_mus.*
34: em_hg_pin.*
35: em_hg_rtd.*
36: em_hg_mam.*
37: em_hg_vrt.*
38: em_sy.*
39: em_hggo_hum.*
40: em_hggo_mus.*
41: em_hggo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------------------|
| 1 | 948 | 100.0 | 948 | 6 | AX087885 Sequence |
| 2 | 948 | 100.0 | 1070 | 6 | AX087884 Sequence |
| 3 | 946.4 | 99.8 | 183455 | 2 | AC008620 Homo sapi |
| 4 | 935.4 | 98.7 | 1192 | 6 | AB065465 Homo sapi |
| 5 | 908.8 | 95.9 | 975 | 6 | AX241794 Sequence |
| 6 | 812 | 85.7 | 236913 | 2 | AL161615 Homo sapi |
| 7 | 810.4 | 85.5 | 948 | 6 | AX375246 Sequence |
| 8 | 810.4 | 85.5 | 1348 | 6 | AX448545 Sequence |
| 9 | 810.4 | 85.5 | 183455 | 2 | AB065675 Homo sapi |
| 10 | 809 | 85.3 | 945 | 6 | AC008620 Homo sapi |
| 11 | 807.4 | 85.2 | 945 | 6 | AX241793 Sequence |
| 12 | 749.6 | 79.1 | 106325 | 2 | AC103024 Homo sapi |
| 13 | 743.2 | 78.4 | 948 | 10 | AY073214 Mus muscu |
| 14 | 743.2 | 78.4 | 123115 | 10 | AL645688 Mouse DNA |
| 15 | 743.2 | 78.4 | 170949 | 2 | AL669850 Mus muscu |
| 16 | 743.2 | 78.4 | 246828 | 2 | AC109931 Rattus no |
| 17 | 679.2 | 71.6 | 246828 | 2 | AY073756 Mus muscu |
| 18 | 674.4 | 71.1 | 981 | 10 | AL645688 Mouse DNA |
| 19 | 674.4 | 71.1 | 123115 | 10 | AL645688 Mouse DNA |
| 20 | 674.4 | 71.1 | 170949 | 2 | AL669850 Mus muscu |
| 21 | 672.4 | 70.9 | 965 | 6 | AX305134 Sequence |
| 22 | 610.8 | 64.4 | 647 | 9 | AF399462 Homo sapi |
| 23 | 552.4 | 58.3 | 650 | 10 | AF399614 Homo sapi |
| 24 | 526.6 | 55.5 | 669 | 10 | AF102533 Mus muscu |
| 25 | 526.6 | 55.5 | 669 | 10 | AF102537 Mus muscu |
| 26 | 525 | 55.4 | 669 | 10 | AF102535 Mus muscu |
| 27 | 442.2 | 46.6 | 954 | 10 | AY073213 Mus muscu |
| 28 | 399 | 42.1 | 82315 | 9 | AL592313 Human DNA |
| 29 | 390.2 | 41.2 | 488 | 6 | AX181424 Sequence |
| 30 | 390.2 | 41.2 | 488 | 6 | AX181424 Sequence |
| 31 | 389.4 | 41.1 | 939 | 6 | AF179751 Gorilla g |
| 32 | 389.4 | 41.1 | 1339 | 9 | AB065952 Homo sapi |
| 33 | 389.4 | 41.1 | 159867 | 9 | AL450303 Human DNA |
| 34 | 386.4 | 40.8 | 159726 | 2 | AC127949 Rattus no |
| 35 | 386.2 | 40.7 | 930 | 10 | AY073559 Mus muscu |
| 36 | 386.2 | 40.7 | 936 | 6 | AX448867 Sequence |
| 37 | 386.2 | 40.7 | 1336 | 6 | AB065953 Homo sapi |
| 38 | 386.2 | 40.7 | 170393 | 10 | AC109605 Mus muscu |
| 39 | 384.6 | 40.6 | 73038 | 2 | AC020884_3 Continuation (4 of |
| 40 | 384.6 | 40.6 | 194086 | 2 | AL662903 Mus muscu |
| 41 | 383 | 40.4 | 184692 | 2 | AC094035 Rattus no |
| 42 | 383 | 40.4 | 195912 | 2 | AC103465 Rattus no |
| 43 | 379.8 | 40.1 | 939 | 6 | AX448873 Sequence |
| 44 | 379.8 | 40.1 | 1339 | 9 | AB065954 Homo sapi |
| 45 | 379.8 | 40.1 | 159867 | 9 | AL450303 Human DNA |

ALIGNMENTS

| RESULT 1 | AX087885 | 948 bp | DNA | linear | PAT 17-MAR-2001 |
|---|---|-----------------------------------|-----|--------|-----------------|
| LOCUS | AX087885 | Sequence 1 from Patent WO0114554. | | | |
| DEFINITION | AX087885 | | | | |
| ACCESSION | AX087885 | | | | |
| VERSION | AX087885.1 | GI:13396877 | | | |
| KEYWORDS | | | | | |
| SOURCE | human. | | | | |
| ORGANISM | Homo sapiens | | | | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | |
| Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | | |
| REFERENCE | 1 (bases 1 to 948) | | | | |
| AUTHORS | Vogeli, G. and Wood, L.S. | | | | |
| TITLE | G protein-coupled receptor expressed in brain | | | | |
| JOURNAL | Patent: WO 0114554-A 1 01-MAR-2001; | | | | |

| FEATURES | Pharmacia & Upjohn Company (US) |
|---------------------------|---|
| SOURCE | Location/Qualifiers |
| CDS | 1. .948 /organism="Homo sapiens" /db_xref="taxon:9606" 1. .948 /note="unnamed protein product" /codon_start=1 /protein_id="CAC34662.1" /db_xref="GI:13396878" /translation="MGRMVNVSYTDGFLLGIFSHSOTDLVLSAVWVFTVALCGNW LFLFLITLDAGLHPMTFFLSQSLMDMLMCNTPPKMAAFLESGKRSISFVGGLD GFEFVLSVSGEGLLGLMAIDYHVAVSPHLPILIMNORVCLQITGSSWARI IDGVQ MVAAMGAPYCGSRSDVHDFCEVOALILKACADPSLFTPLLFCACFMLLPFSITMAS YACFLGAVLIRIRSRQAMKMLATCSHLLATWLEFYGAEMWYLRPRRYRAPSBDKVAS IFVFLVPLNPLPLYSLRNKGVMALSKGLDRCRIGSOH" |
| BASE COUNT | 172 a 278 c 230 g 268 t |
| ORIGIN | |
| Query Match | 100.0%; Score 948; DB 6; Length 948; |
| Best Local Similarity | 100.0%; Pred. No. 8.6e-242; |
| Matches 948; Conservative | 0; Mismatches 0; Indels 0; Gaps 0; |
| OY | 1 ATGGGAAGATGGGTCACACCAAGCTCCACACACATAGGCTCTTCTCTTGCGCATCTTTCC 60 |
| Db | 1 ATGGGAAGATGGGTCACACCAAGCTCCACACAGATGGCTCTTCTCTTGCGCATCTTTCC 60 |
| OY | 61 CACAGCCAGACTGCACCTTGCTCTTCTCTGCACTTAATGAGTGCTTACAGTGGCCCTC 120 |
| Db | 61 CACAGCCAGACTGCACCTTGCTCTTCTCTGCACTTAATGAGTGCTTACAGTGGCCCTC 120 |
| OY | 121 TGTGGGAATGTCTCTCTATCTTCTCATCTACCTGACAGCTGGACCTCACACCCCATG 180 |
| Db | 121 TGTGGGAATGTCTCTCTATCTTCTCATCTACCTGACAGCTGGACCTCACACCCCATG 180 |
| OY | 181 TACTCTTCTCTCAGCAGCAGCTCCCTCATATGAGACCTCATATGGTGCTGTAACTTGGCA 240 |
| Db | 181 TACTCTTCTCTCAGCAGCAGCTCCCTCATATGAGACCTCATATGGTGCTGTAACTTGGCA 240 |
| OY | 241 AAGATGAGCAGCCAACTCTCTGTCTGCGAGAGATCCATCTCTTTGTGGGCTGTGGCATA 300 |
| Db | 241 AAGATGAGCAGCCAACTCTCTGTCTGCGAGAGATCCATCTCTTTGTGGGCTGTGGCATA 300 |
| OY | 301 CAAATTTGGCTTTTGTCTCTCTTGTGGGAGTCTGAGGGGCTTCTCTGGACTCATAGGCT 360 |
| Db | 301 CAAATTTGGCTTTTGTCTCTCTTGTGGGAGTCTGAGGGGCTTCTCTGGACTCATAGGCT 360 |
| OY | 361 TATGACCACTACGTGGCCGTTAGCCACCCACTTCACTATCCCATCCCTCATGAAATCAGAG 420 |
| Db | 361 TATGACCACTACGTGGCCGTTAGCCACCCACTTCACTATCCCATCCCTCATGAAATCAGAG 420 |
| OY | 421 GTCCTGTCTCCAGATTACTGAGGAGCTCTGGGCGCTTTGGGATTAATAGATGAGTATAG 480 |
| Db | 421 GTCCTGTCTCCAGATTACTGAGGAGCTCTGGGCGCTTTGGGATTAATAGATGAGTATAG 480 |
| OY | 481 AATGGGAGCAGCATGGGCTTACCTTACTAGTGAGCTTAAGAGAGGTGATCATTCTTCTGT 540 |
| Db | 481 AATGGGAGCAGCATGGGCTTACCTTACTAGTGAGCTTAAGAGAGGTGATCATTCTTCTGT 540 |
| OY | 541 GAGGTACAAAGCTTTAATGAAGCTGGCGTGTGAGACACTTCCCTTTTGTACACCCCTCTC 600 |
| Db | 541 GAGGTACAAAGCTTTAATGAAGCTGGCGTGTGAGACACTTCCCTTTTGTACACCCCTCTC 600 |
| OY | 601 TTTGGTCTCTGTCTGTATAGTCTTCTCTCTCTCTCTCATATCATATAGGCGCTCATAGCT 660 |
| Db | 601 TTTGGTCTCTGTCTGTATAGTCTTCTCTCTCTCTCTCATATCATATAGGCGCTCATAGCT 660 |
| OY | 661 TGCATCTACAGGGGCTGTCTCCGGAATACGCTCTGCTCAGGCGCTGGAAAAAGCCCTGGCC 720 |
| Db | 661 TGCATCTACAGGGGCTGTCTCCGGAATACGCTCTGCTCAGGCGCTGGAAAAAGCCCTGGCC 720 |
| OY | 721 AACTGCTCTCTCCACCTTAACAGCTGTACACCTCTTCTAATGGGAGCCATGTTCATGTAC 780 |
| Db | 721 AACTGCTCTCTCCACCTTAACAGCTGTACACCTCTTCTAATGGGAGCCATGTTCATGTAC 780 |

| Db | 721 | ACBCTCCTCCACCTTAACAGCTGTCAACCCCTCTCTATGAGGGGACGCCATGTCATGTAC | 780 |
|-----------------------|--|---|----------------------------|
| Qy | 781 | CTGAGGCGCTAGGCGGTACCGGGGCGCCCTTGACCATGACAAAGGTGGCGCTCATATCTTACACA | 840 |
| Db | 781 | CTGAGGCGCTAGGCGGTACCGGGGCGCCCTTGACCATGACAAAGGTGGCGCTCATATCTTACACA | 840 |
| Qy | 841 | GTCCTTACTCCCATGCTGAACCCCTCATTTTACACCTTGAGGAATGGGAGGTGATGGG | 900 |
| Db | 841 | GTCCTTACTCCCATGCTGAACCCCTCATTTTACACCTTGAGGAATGGGAGGTGATGGG | 900 |
| Qy | 901 | GCACCTAGGAAGGGGCTGACCGCTGCAGAGATGTGGACGCACACTGCA | 948 |
| Db | 901 | GCACCTAGGAAGGGGCTGACCGCTGCAGAGATGTGGACGCACACTGCA | 948 |
| RESULT 2 | | | |
| LOCUS | AX087894 | 1070 bp | DNA linear PAT 17-MAR-2001 |
| DEFINITION | Sequence 10 from Patent W00114554. | | |
| ACCESSION | AX087894 | | |
| VERSION | AX087894.1 | GI:13396883 | |
| KEYWORDS | human. | | |
| SOURCE | ORGANISM | Homo sapiens | |
| | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| | | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. | |
| REFERENCE | 1 (bases 1 to 1070) | | |
| AUTHORS | Vogeli, G. and Wood, L.S. | | |
| TITLE | G protein-coupled receptor expressed in brain | | |
| JOURNAL | Patent: WO 0114554-A 10-01-MAR-2001; Pharmacia & Upjohn Company (US) | | |
| FEATURES | Location/Qualifiers | | |
| Source | 1..1070 | | |
| | /organism="Homo sapiens" | | |
| | /db_xref="taxon:9606" | | |
| BASE COUNT | 207 a 305 c 254 g 304 t | | |
| ORIGIN | | | |
| Query Match | 100.0% | Score 948: | DB 6: Length 1070: |
| Best Local Similarity | 100.0% | Pred. No. 8.6e-242: | |
| Matches 948: | Conservative 0: Mismatches 0: Indels 0: Gaps 0: | | |
| Qy | 1 | ATGGGAAGATGGGTGAACACAGTCTACACAGATGGGCTTCTCTCTGAGGATCTTTTCC | 60 |
| Db | 123 | ATGGGAAGATGGGTGAACACAGTCTACACAGATGGGCTTCTCTCTGAGGATCTTTTCC | 182 |
| Qy | 61 | CACAGCCAGACTGACCTTGTCTCTTCTGTGCAGTTATGTGTGCTTCACAGTGGCCTC | 120 |
| Db | 183 | CACAGCCAGACTGACCTTGTCTCTTCTGTGCAGTTATGTGTGCTTCACAGTGGCCTC | 242 |
| Qy | 121 | TGTGGAGATGTCCTCATCTTCTCATCTACATCACTGAGAGCGTGGACCTTCACACCCCATG | 180 |
| Db | 243 | TGTGGAGATGTCCTCATCTTCTCATCTACATCACTGAGAGCGTGGACCTTCACACCCCATG | 302 |
| Qy | 181 | TACTCTTCTCAGCCAGCTCTCCCTCATGACCTCATGTTGGTCTGTAACTTGTGCCA | 240 |
| Db | 303 | TACTCTTCTCAGCCAGCTCTCCCTCATGACCTCATGTTGGTCTGTAACTTGTGCCA | 362 |
| Qy | 241 | AAGATGGCAGCCACTTCTCTGTGTGGCAGAGAGTCCATCTCTTTGTGGGCTGTGGCATA | 300 |
| Db | 363 | AAGATGGCAGCCACTTCTCTGTGTGGCAGAGAGTCCATCTCTTTGTGGGCTGTGGCATA | 422 |
| Qy | 301 | CAAAATGGCTTTTGTCTCTCTTGTGGGAGTGTGAGGGGCTTGTCTGGAGATCATGGCT | 360 |
| Db | 423 | CAAAATGGCTTTTGTCTCTCTTGTGGGAGTGTGAGGGGCTTGTCTGGAGATCATGGCT | 482 |
| Qy | 361 | TATGACCACTACGTGGCGTTAGCCACCCACTTCACTATCCCATCCTCATGAATCAGAGG | 420 |
| Db | 483 | TATGACCACTACGTGGCGTTAGCCACCCACTTCACTATCCCATCCTCATGAATCAGAGG | 542 |
| Qy | 421 | GTCCTCTCCAGATTACTGGGAGCTCTGTGGCGCTTTGGATATATAGATGAGTGAATTGAG | 480 |
| Db | 543 | GTCCTCTCCAGATTACTGGGAGCTCTGTGGCGCTTTGGATATATAGATGAGTGAATTGAG | 602 |

| | | | |
|----|------|---|------|
| QY | 481 | ATGTGGAGGCATGGGCTTACCTTACCTGATGGCTCAAGAGGCTGGATCACATTTTTCGT | 540 |
| Db | 603 | ATGGTGGAGGCATGGGCTTACCTTACCTGATGGCTCAAGAGGCTGGATCACATTTTTCGT | 662 |
| QY | 541 | GAGTCAACGCTTTATTTAGAGCTGGCGTGTGAGACACTTCCCTTTTGTGACACCTCTGC | 600 |
| Db | 663 | GAGTCAACGCTTTATTTAGAGCTGGCGTGTGAGACACTTCCCTTTTGTGACACCTCTGC | 722 |
| QY | 601 | TTTGCTTACCTGTCGTCCTGATGCTTCTCTCCCTCTCATCATATGAGCCCTCATGCT | 660 |
| Db | 723 | TTTGCTTACCTGTCGTCCTGATGCTTCTCTCCCTCTCATCATATGAGCCCTCATGCT | 782 |
| QY | 661 | TGCATCTTAGGGGCGCTGTGCTCCGAATACGCTGTGCTGACGCGCTGGAAAAAACCCCTGGCC | 720 |
| Db | 783 | TGCATCTTAGGGGCGCTGTGCTCCGAATACGCTGTGCTGACGCGCTGGAAAAAACCCCTGGCC | 842 |
| QY | 721 | ACCGCTCCTCCCAACCTACAGCTGTACCCCTCTCTATGAGGGGACGCATGTTCAATGTAC | 780 |
| Db | 843 | ACCGCTCCTCCCAACCTACAGCTGTGTACCCCTCTCTATGAGGGGACGCATGTTCAATGTAC | 902 |
| QY | 781 | CTGAGGCGCTAGGCGCTACCGGGGCCCTTGAGCCATGACAAGAGTGCGCCCTCATCTTCTACACA | 840 |
| Db | 903 | CTGAGGCGCTAGGCGCTACCGGGGCCCTTGAGCCATGACAAGAGTGCGCCCTCATCTTCTACACA | 962 |
| QY | 841 | GTCCTTACTCCCATGCTGAACCCCTCATTTTACAGCTTGAGGAATGGGAGGTGATGGGG | 900 |
| Db | 963 | GTCCTTACTCCCATGCTGAACCCCTCATTTTACAGCTTGAGGAATGGGAGGTGATGGGG | 1022 |
| QY | 901 | GCACCTGAGAGGGGCGTGGACCGCTGCAGAGGATTTGGACACCCAGCACTGA | 948 |
| Db | 1023 | GCACCTGAGAGGGGCGTGGACCGCTGCAGAGGATTTGGACACCCAGCACTGA | 1070 |

```

RESULT 3
AC008620
LOCUS
DEFINITION Homo sapiens chromosome 5 clone CTB-14A14, WORKING DRAFT SEQUENCE,
8 ordered pieces.
AC008620
AC008620.8 GI:19224690
VERSION
KEYWORDS HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE
ORGANISM Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 183455)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 183455)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-MAR-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 183455)
DOE Joint Genome Institute.
Direct Submission
Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Mar 7, 2002 this sequence version replaced gi:13676956.
-----Genome Center
Center: Joint Genome Institute
Center code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 69594
Center clone name: CIT978SKB_14A14
-----
Summary Statistics
Consensus bases at least Q40
Consensus quality: 180052 bases at least Q30

```

```

Consensus quality: 181633 bases at least Q20
Estimated insert size: 197550; agarose-fp estimation
Estimated insert size: 182755; sum-of-contigs estimation
Quality coverage: 6.16 in Q20 bases; agarose-fp estimation
Quality coverage: 6.66 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1      8540: contig of 8540 bp in length
*      8541      8640: gap of unknown length
*      8641      37607: contig of 28967 bp in length
*      37608      37707: gap of unknown length
*      37708      39853: contig of 2146 bp in length
*      39854      39953: gap of unknown length
*      39954      109099: contig of 69146 bp in length
*      109100      109199: gap of unknown length
*      109200      113256: contig of 4057 bp in length
*      113257      113356: gap of unknown length
*      113357      116278: contig of 2922 bp in length
*      116279      116378: gap of unknown length
*      116379      160370: contig of 43992 bp in length
*      160371      160470: gap of unknown length
*      160471      183455: contig of 22985 bp in length.
Location/Qualifiers

```

| BASE COUNT | 47989 a | 44173 c | 44560 g | 45962 t | 771 others |
|---------------------------|--|---------------|-----------|---------|------------|
| ORIGIN | /clone="CTB-14A14" /clone_lib="caltech human BAC library B" | | | | |
| Query Match | 99.8% Score 946.4; DB 2; Length 183455; | | | | |
| Best Local Similarity | 99.9% Pred. No. 2.4e-241; | | | | |
| Matches 947; Conservative | 0; | Mismatches 1; | Indels 0; | Gaps 0; | |
| QY 1 | ATGGAGATGGGTGAACCACTCCACAGATGGCTTCCTCTTGGGCATCTTTTCC | 60 | | | |
| Db 102197 | ATGGAGATGGGTGAACCACTCCACAGATGGCTTCCTCTTGGGCATCTTTTCC | 102256 | | | |
| QY 61 | CACAGCCAGACTGACCTTGTCTCTTCTGTCAGTTATGTTGCTTTACACAGTGGCCCTC | 120 | | | |
| Db 102257 | CACAGCCAGACTGACCTTGTCTCTTCTGTCAGTTATGTTGCTTTACACAGTGGCCCTC | 102316 | | | |
| QY 121 | TGTGGGATGTCCTCCTCATCTCTCTCATTCACCGGAGGCGGAGCTTCACACCCCATG | 180 | | | |
| Db 102317 | TGTGGGATGTCCTCCTCATCTCTCTCATTCACCGGAGGCGGAGCTTCACACCCCATG | 102376 | | | |
| QY 181 | TACTTCTTCCACAGCACTCCCTCATGAGACCTCATGTTGGTGTGAACATTGTGCCA | 240 | | | |
| Db 102377 | TACTTCTTCCACAGCACTCCCTCATGAGACCTCATGTTGGTGTGAACATTGTGCCA | 102436 | | | |
| QY 241 | AAGATGCGAGCCAACTTCCTGTGCGAGAGATCCATCTCTTGTGGGCTGTGGCATA | 300 | | | |
| Db 102437 | AAGATGCGAGCCAACTTCCTGTGCGAGAGATCCATCTCTTGTGGGCTGTGGCATA | 102496 | | | |
| QY 301 | CAATTTGGCTTTTGTCTCTCTGTGGGATCTGAGGGGCTCTTCTGTGGACTATGGCT | 360 | | | |
| Db 102497 | CAATTTGGCTTTTGTCTCTCTGTGGGATCTGAGGGGCTCTTCTGTGGACTATGGCT | 102556 | | | |
| QY 361 | TATAGCACTACGTGGCGGTAGCACCCACTCTCATATCCATCCCTATCAATAGAGG | 420 | | | |
| Db 102557 | TATAGCGGTACGTGGCGGTAGCACCCACTCTCATATCCATCCCTATCAATAGAGG | 102616 | | | |
| QY 421 | GTCGTCTCCAGATTACTGGAGCTCTGGGCTTTGGGATTAATAGATGAGTGATTTAG | 480 | | | |

| | | | | |
|------------|---|---|---|------------------------|
| | Db | 102617 | GTCGTGCTCCAGATTACTGGGAGCGTCTGGGCCCTTTGGGTAATAGATGAGTGATTCAG | 102616 |
| OY | 481 | ATGGTGGCACCACCATGGGCTTAACCTTAAGCTGAAGGACCGTGATCACTTTTTCTGT | 540 | |
| Db | 102677 | ATGGTGGCACCCATGGGCTTAACCTTAAGCTGAAGGACCGTGATCACTTTTTCTGT | 102736 | |
| OY | 541 | GAGGTACAAGCTTATTGAAGCTGGCGCTGGCAGACACTTCCCTTTTGGACACCTCTGC | 600 | |
| Db | 102737 | GAGGTACAAGCTTATTGAAGCTGGCGCTGGCAGACACTTCCCTTTTGGACACCTCTGC | 102796 | |
| OY | 601 | TTTTCCTGCTGTGTCTTCATGCTTCTCTCTCCCTTCCTTCATCATCATGCGCTCTATGCT | 660 | |
| Db | 102797 | TTTTCCTGCTGTGTCTTCATGCTTCTCTCTCCCTTCCTTCATCATCATGCGCTCTATGCT | 102856 | |
| OY | 661 | TGCATCTTAGGGGCGTGAGCTCCCGAATACGCGTCCCTCAGGCGTGGAAAAAACCCCTGGCC | 720 | |
| Db | 102857 | TGCATCTTAGGGGCGTGAGCTCCCGAATACGCGTCCCTCAGGCGTGGAAAAAACCCCTGGCC | 102916 | |
| OY | 721 | ACCTGCTCTCTCCACCTAACAGCTGATACCCCTCTTCTATATGGGAGCAGCATGTTGATGTAC | 780 | |
| Db | 102917 | ACCTGCTCTCTCCACCTAACAGCTGATACCCCTCTTCTATATGGGAGCAGCATGTTGATGTAC | 102976 | |
| OY | 781 | CTGAGGCGCTAAGCGCTACCGGGCCCCCTAAGCATATACAAGTGGCGCTCATATCTTACACA | 840 | |
| Db | 102977 | CTGAGGCGCTAAGCGCTACCGGGCCCCCTAAGCATATACAAGTGGCGCTCATATCTTACACA | 103036 | |
| OY | 841 | GTCCCTTAATCCCATGCGAAACCCCTCATTTACAGCTTTGAGAATGGGAGATGATGGGG | 900 | |
| Db | 103037 | GTCCCTTAATCCCATGCGAAACCCCTCATTTACAGCTTTGAGAATGGGAGATGATGGGG | 103096 | |
| OY | 901 | GCACCTGAGAGANGGGGCTGGACCCGCTGACAGATTGGGACACCGACGACTGA | 948 | |
| Db | 103097 | GCACCTGAGAGANGGGGCTGGACCCGCTGACAGATTGGGACACCGACGACTGA | 103144 | |
| RESULT 4 | | | | |
| LOCUS | AB065465 | 1192 bp | DNA | Linear PRI 23-Jul-2002 |
| DEFINITION | Homo sapiens gene for seven transmembrane helix receptor, complete cds, isolate:CBRC7TM_28. | | | |
| ACCESSION | AB065465 | | | |
| VERSION | AB065465.1 | GI:21928240 | | |
| KEYWORDS | | | | |
| SOURCE | | | | |
| ORGANISM | Homo sapiens (isolate:CBRC7TM_28) DNA. | | | |
| REFERENCE | Homo sapiens | | | |
| AUTHORS | Euhayrolta; Metazora; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| TITLE | 1 Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S., Tsutsumi,S., Aburatani,H., Asai,K. and Akiyama,Y. Genome-wide discovery and analysis of human seven transmembrane helix receptor genes Unpublished 2 (bases 1 to 1192) Suwa,M. | | | |
| JOURNAL | Direct Submission | | | |
| REFERENCE | Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST); 2-41-6 Aoml Koto-ku, Tokyo 135-0064, Japan (E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/, Tel:81-3-3599-8080, Fax:81-3-3599-8081) | | | |
| AUTHORS | This sequence is a seven transmembrane helix receptor candidate predicted from the whole human genome sequences using our automated system that contains programs of gene finding(GeneDecoder), sequence search, motif-domain assignment and transmembrane helix prediction. | | | |
| JOURNAL | And the sequence is submitted by the collaborative project between [Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST)] and [Genome Science Division, Research Center for Advanced Science and Technology (RCAS)], University of Tokyo]. | | | |
| COMMENT | | | | |
| FEATURES | | | | |
| Source | Location/Qualifiers | | | |
| | 1..1192 | | | |

| | | | |
|-----------------------|--------------|--|----------------------|
| | | /organism="Homo sapiens" | |
| | | /isolate="CBRC7TM_28" | |
| | | /db_xref="taxon:9606" | |
| | | /chromosome="5" | |
| CDS | | 201..992 | |
| | | /evidence-not_experimental | |
| | | /product="seven transmembrane helix receptor" | |
| | | /protein_id="BAC05724.1" | |
| | | /db_xref="GI:21928241" | |
| | | /translation="MGRWVNDSTYDFEFLLGITSHSDTLVLSAFNVFVALGCVN LLFLVLIDAGLHPMYEFLSQLSLMDLMVCNTVPKMANFLSGRSISIVGGIGID GFPSVLSEGLGLMALADRYVAIVASPLRPILMNDRVCLQITSSMARQIDGVLD MVAMGILPYCGSRVDHFPEVOVALHLACADTSLPDTLFAFCVFMLILPFSIIMS YALSRCAENLTCSGLEKSPGHILLPRNSCHPLMSHVHPA" | |
| BASE COUNT | 232 a | 339 c | 285 g 336 t |
| ORIGIN | | | |
| Query Match | 98.7% | Score 935.4: | DB 9; Length 1192; |
| Best Local Similarity | 99.8%; | Pred. No. 2e-238; | |
| Matches 947; | Conservative | 0; Mismatches | 1; Indels 1; Gaps 1; |
| Oy | 1 | ATGGGAAGATGGGTGTAACAGCTCCTACTACACAGATGGCTTCCTCTTGGGCATCTTTCC | 60 |
| Db | 201 | ATGGGAAGATGGGTGTAACAGAGCTCTACACAGATGGCTTCCTCTTGGGCATCTTTCC | 260 |
| Oy | 61 | CACAGCCAGACTGACCTGTCTCTCTCTGCAAGTAAAGTGGTCTCACACATGGCCCCC | 120 |
| Db | 261 | CACAGCCAGACTGACCTGTCTCTCTCTGCAAGTAAAGTGGTCTCACACATGGCCCCC | 320 |
| Oy | 121 | TGTGGGAATGTCTCTCTCTCTCTCTCATCTTCACATGAGAGCTGAGATTACACCCCAGT | 180 |
| Db | 321 | TGTGGGAATGTCTCTCTCTCTCTCTCATCTTCACATGAGAGCTGAGATTACACCCCAGT | 380 |
| Oy | 181 | TACTTCTCTCTCAGCCAGCTCTCCCTCATGACCTCATATGTTGGTGTAACTTTGGCCA | 240 |
| Db | 381 | TACTTCTCTCTCAGCCAGCTCTCCCTCATGACCTCATATGTTGGTGTAACTTTGGCCA | 440 |
| Oy | 241 | AAGATGGCAAGCCAACTTCCTGTCTGGAGAGAAGTCCATCTCTTTGGGGCTGGGCATA | 300 |
| Db | 441 | AAGATGGCAAGCCAACTTCCTGTCTGGAGAGAAGTCCATCTCTTTGGGGCTGGGCATA | 500 |
| Oy | 301 | CAAAATTGGCTTTTTGTCTCTCTTGTGGAGTCTGAGGGGCTCTGTGCGGACTATGGCT | 360 |
| Db | 501 | CAAAATTGGCTTTTTGTCTCTCTTGTGGAGTCTGAGGGGCTCTGTGCGGACTATGGCT | 560 |
| Oy | 361 | TATGACCACCTACGTGGCCCTTAGCCACCACCTTCACTATCCCATCTCATGAAATGAGAG | 420 |
| Db | 561 | TATGACCACCTACGTGGCCCTTAGCCACCACCTTCACTATCCCATCTCATGAAATGAGAG | 620 |
| Oy | 421 | GTCCTGCTCCAACTTACCTGAGAGCTCCTGGGCCCTTTGGGATTAATAGATGAGAGATTAG | 480 |
| Db | 621 | GTCCTGCTCCAACTTACCTGAGAGCTCCTGGGCCCTTTGGGATTAATAGATGAGAGATTAG | 680 |
| Oy | 481 | ATGGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGGTGATCACTTTTCTGT | 540 |
| Db | 681 | ATGGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGGTGATCACTTTTCTGT | 740 |
| Oy | 541 | GAGGTACAAGCTTTATGAAGCTGGCCTGTGACAGACACTTCCCTTTTGAACCCCTCTC | 600 |
| Db | 741 | GAGGTACAAGCTTTATGAAGCTGGCCTGTGACAGACACTTCCCTTTTGAACCCCTCTC | 800 |
| Oy | 601 | TTTGTGCTGTGTCTTCAANGCTTCCTCCCTTCATCATCATATGAGCGTCTCATGTGT | 660 |
| Db | 801 | TTTGTGCTGTGTCTTCAANGCTTCCTCCCTTCATCATCATATGAGCGTCTCATGTGT | 860 |
| Oy | 661 | TGCATC-TTAGGGGCGTCTCCGGAATAGCTTCTGCTACAGGCCCTGGAAAAAACCCCTGGC | 719 |
| Db | 861 | TGCATCTTAGGGGCGTCTGCGAATAGCTTCTGCTACAGGCCCTGGAAAAAACCCCTGGC | 920 |
| Oy | 720 | CACCTGCTCTCCACACTAAACAGCTGTCAACCTCTTCTATAGGGGCAAGCCATGTTCAATGA | 779 |
| Db | 921 | CACCTGCTCTCCACACTAAACAGCTGTCAACCTCTTCTATAGGGGCAAGCCATGTTCAATGA | 980 |

| QY | 780 | CCTGAGGCCCTAGGCCCTACCGGGCCCTGAGCCATGACAGATGGCCCTATCTTACAC | 839 |
|-----------------------|--|---|----------------------------|
| Db | 981 | CCTGAGGCCCTAGGCCCTACCGGGCCCTGAGCCATGACAGATGGCCCTATCTTACAC | 1040 |
| QY | 840 | AGTCTTACTCCCATGCTGAAACCCCTCATTTTACAGCTTGAGAAATGGGAGGTGATGG | 899 |
| Db | 1041 | AGTCTTACTCCCATGCTGAAACCCCTCATTTTACAGCTTGAGAAATGGGAGGTGATGG | 1100 |
| QY | 900 | GGCACTGAGGAAGGGGCTGACCGCTGCAGAGATTGGCAGCCAGCACTGA | 948 |
| Db | 1101 | GGCACTGAGGAAGGGGCTGACCGCTGCAGAGATTGGCAGCCAGCACTGA | 1149 |
| RESULT 5 | | | |
| LOCUS | AX241794 | 975 bp | DNA linear PAT 26-SEP-2001 |
| DEFINITION | Sequence 542 from Patent WO0127158. | | |
| ACCESSION | AX241794 | | |
| VERSION | AX241794.1 | GI:15798669 | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | Olfactory receptor sequences | | |
| JOURNAL | Patent: WO 0127158-A 542 19-APR-2001; | | |
| FEATURES | | | |
| source | | | |
| BASE COUNT | 174 a 290 c 237 g 274 t | | |
| ORIGIN | | | |
| Query Match | 95.9%; Score 908.8; DB 6; Length 975; | | |
| Best Local Similarity | 99.5%; Pred. No. 2.4e-231; | | |
| Matches | 943; Conservative 0; Mismatches 2; Indels 3; Gaps 3. | | |
| QY | 1 | ATGGAGAGTGGGTGAGACCACTCTACACAGATGGCTTCTTCTGAGCATCTTTTCC | 60 |
| Db | 1 | ATGGAGAGTGGGTGAGACCACTCTACACAGATGGCTTCTTCTGAGCATCTTTTCC | 60 |
| QY | 61 | CACACCAGACTGACCTGTCTCTCTCTCTGACGTTAATGATGCTTTACAGTGGCCCTC | 120 |
| Db | 61 | CACACCAGACTGACCTGTCTCTCTCTCTGACGTTAATGATGCTTTACAGTGGCCCTC | 120 |
| QY | 121 | TGTGGAGATGTCCTCTCATCTCTCTCATCTTCACTGAGAGCGTGCACACCCCATG | 180 |
| Db | 121 | TGTGGAGATGTCCTCTCATCTCTCTCATCTTCACTGAGAGCGTGCACACCCCATG | 180 |
| QY | 181 | TACTTCTTCTCAGCAGACTCTCCCTCATGACGTCATGATGATGATGCTGTAACTTGTGCA | 240 |
| Db | 181 | TACTTCTTCTCAGCAGACTCTCCCTCATGACGTCATGATGATGATGCTGTAACTTGTGCA | 240 |
| QY | 241 | AAGATGAGCAGCACTTCTCTGTGTGACAGAACTCATCTCTTTGTGGGCTGTGGCATA | 300 |
| Db | 241 | AAGAT- GA GAGCAACTTCTCTGTGTGCA- GA AGTCCATCTCTTTGTGGGCTGTGGCATA | 298 |
| QY | 301 | CAAAATGGCTTTTGTCTCTCTTGTGGAGATCGAGGGGCTCTGCTGGACATAGGCT | 360 |
| Db | 299 | CAAAATGGCTTTTGTCTCTCTTGTGGAGATCGAGGGGCTCTGCTGGACATAGGCT | 358 |
| QY | 361 | TATGACCACTACGTCGCGTTAGCCACCCACATTCATATCCATCCCTATGATGATAGAG | 420 |
| Db | 359 | TATGACCCCTACGTCGCGTTAGCCACCCACATTCATATCCATCCCTATGATGATAGAG | 418 |
| QY | 421 | GTCGTCTTCACGATTAACGGAGGCTCTCGGGCTTTGGGATTAATGATGAGTGAATTAG | 480 |
| Db | 419 | GTCGTCTTCACGATTAACGGAGGCTCTCGGGCTTTGGGATTAATGATGAGTGAATTAG | 478 |

| | | | |
|----|-----|---|-----|
| OY | 481 | ATGAGGAGGCAATGGGCTACCTTACGTGAGGCTGAAGGAGGTGATACCTTTTCTGT | 540 |
| Dd | 479 | ATGGTGGAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGGTGATACCTTTTCTGT | 538 |
| OY | 541 | GAGGTACAAAGCTTATTTGAAGCTGGGCTGTGCAGACACTTCCCTTTTGGACACCCTCTC | 600 |
| Dd | 539 | GAGGTACAAAGCTTATTTGAAGCTGGGCTGTGCAGACACTTCCCTTTTGGACACCCTCTC | 598 |
| OY | 601 | TTTTCCTTGCTGTCTTCCTCAGCTTTCCTTCCTTCCTTCATATATAGGCTCCTATGCT | 660 |
| Dd | 599 | TTTTCCTTGCTGTCTTCCTCAGCTTTCCTTCCTTCCTTCATATATAGGCTCCTATGCT | 658 |
| OY | 661 | TGCATCCTAGGAGGCTGTGCTCCCAATACGCTTCGCTCAGAGGCTGGAAAAAGCCCTGGCC | 720 |
| Dd | 659 | TGCATCCTAGGAGGCTGTGCTCCCAATACGCTTCGCTCAGAGGCTGGAAAAAGCCCTGGCC | 718 |
| OY | 721 | ACCTGCTCTCTCCACCTAACAGCTGTACCCCTCTCTATAGGGGACGCCATGTCATGTAC | 780 |
| Dd | 719 | ACCTGCTCTCTCCACCTAACAGCTGTACCCCTCTCTATAGGGGACGCCATGTCATGTAC | 777 |
| OY | 781 | CTGAGGCGCTAGGGGCTACCGGGGCGCCCTAGCCAMTGAAGGTGGCCCTATGCTTCACACA | 840 |
| Dd | 778 | CTGAGGCGCTAGGGGCTACCGGGGCGCCCTAGCCAMTGAAGGTGGCCCTATGCTTCACACA | 837 |
| OY | 841 | GTCCTTACTCCATGCTGANAACCCCTCATTTTACACTTGAGGAATGGGAGGTGATGGGG | 900 |
| Dd | 838 | GTCCTTACTCCATGCTGANAACCCCTCATTTTACAGTTGAGGAATGGGAGGTGATGGGG | 897 |
| OY | 901 | GCACTAGAGAAAGGGGCTGGACCGCTCAGAGTTTGGCAGACAGCACTGA 948 | |
| Dd | 898 | GCACTAGAGAAAGGGGCTGGACCGCTCAGAGTTTGGCAGACAGCACTGA 945 | |

| | | | | |
|------------|--|------------|-----|------------------------|
| RESULT 6 | | | | |
| AL161615/c | | | | |
| LOCUS | AL161615 | 236913 bp | DNA | linear HTG_06-DEC-2001 |
| DEFINITION | Homo sapiens chromosome 13 clone RP11-324K20, *** SEQUENCING IN PROGRESS ***, 52 unordered pieces. | | | |
| ACCESSION | AL161615 | | | |
| VERSION | AL161615.3 | GI:9863602 | | |
| KEYWORDS | HTG; HTGS_PHRASE1; HTGS_CANCELLED. | | | |
| SOURCE | human. | | | |
| ORGANISM | Homo sapiens | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. | | | |
| REFERENCE | 1 | | | |
| AUTHORS | Burton,J. | | | |
| TITLE | Direct Submission | | | |
| JOURNAL | Submitted (05-DEC-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humbug@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Aug 21, 2000 this sequence version replaced gi:7327747. | | | |
| COMMENT | | | | |

```

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
Project Information
Center project name: BA324K20
Summary Statistics
Sequencing program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator; Big Dye; 100% of reads
Consensus quality: 204106 bases at least Q40
Consensus quality: 217692 bases at least Q30
Consensus quality: 225366 bases at least Q20
Insert size: 231813; sum-of-contrigs
Insert size: 232359; 21.7% error; agarose-fp
Quality coverage: 2.25x in Q20 bases; sum-of-contrigs Quality
Coverage: 2.82x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently

```

* consists of 52 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 4992 5091: contig of 4991 bp in length
 5092 7436: contig of 2345 bp in length
 7437 7537: gap of 100 bp
 7537 25274: contig of 17738 bp in length
 25275 25374: gap of 100 bp
 25375 28650: contig of 3276 bp in length
 28651 28750: gap of 100 bp
 28751 31057: contig of 2307 bp in length
 31058 31157: gap of 100 bp
 31158 36458: contig of 5301 bp in length
 36459 36558: gap of 100 bp
 36559 41612: contig of 5054 bp in length
 41613 41712: gap of 100 bp
 41713 44933: contig of 3221 bp in length
 44934 45033: gap of 100 bp
 45034 47911: contig of 2878 bp in length
 47912 48011: gap of 100 bp
 48012 50422: contig of 2411 bp in length
 50423 50522: gap of 100 bp
 50523 54916: contig of 4394 bp in length
 54917 55016: gap of 100 bp
 55017 57305: contig of 2289 bp in length
 57306 57405: gap of 100 bp
 57406 63866: contig of 6461 bp in length
 63867 63966: gap of 100 bp
 63967 67393: contig of 3427 bp in length
 67394 67493: gap of 100 bp
 67494 69948: contig of 2455 bp in length
 69949 70048: gap of 100 bp
 70049 72223: contig of 2174 bp in length
 72223 72322: gap of 100 bp
 72323 74676: contig of 2354 bp in length
 74677 74776: gap of 100 bp
 74777 76906: contig of 2130 bp in length
 76907 77006: gap of 100 bp
 77007 81224: contig of 4218 bp in length
 81225 81324: gap of 100 bp
 81325 85907: contig of 4583 bp in length
 85908 86007: gap of 100 bp
 86008 88623: contig of 2616 bp in length
 88624 88723: gap of 100 bp
 88724 97798: contig of 9075 bp in length
 97799 97898: gap of 100 bp
 97899 103226: contig of 5328 bp in length
 103227 103326: gap of 100 bp
 103327 108117: contig of 4791 bp in length
 108118 108217: gap of 100 bp
 108218 113719: contig of 4502 bp in length
 112720 112819: gap of 100 bp
 112820 132527: contig of 13708 bp in length
 132528 132627: gap of 100 bp
 132628 135076: contig of 2449 bp in length
 135077 135176: gap of 100 bp
 135177 137721: contig of 2545 bp in length
 137722 137821: gap of 100 bp
 137822 13980: contig of 2159 bp in length
 13981 140080: gap of 100 bp
 140081 144461: contig of 4381 bp in length
 144462 144561: gap of 100 bp
 144562 146775: contig of 2214 bp in length
 146776 146875: gap of 100 bp
 146876 150579: contig of 3704 bp in length
 150580 150679: gap of 100 bp
 150680 155630: contig of 4951 bp in length
 155631 155730: gap of 100 bp

FEATURES

source

155731 158259: contig of 2529 bp in length
 158260 158359: gap of 100 bp
 158360 160435: contig of 2076 bp in length
 160436 160535: gap of 100 bp
 160536 166007: contig of 5472 bp in length
 166008 166107: gap of 100 bp
 166108 169792: contig of 3688 bp in length
 169793 169892: gap of 100 bp
 169893 178396: contig of 8504 bp in length
 178397 178496: gap of 100 bp
 178497 182026: contig of 3530 bp in length
 182027 182126: gap of 100 bp
 182127 184776: contig of 2650 bp in length
 184777 184876: gap of 100 bp
 184877 189011: contig of 4135 bp in length
 189012 189111: gap of 100 bp
 189112 196337: contig of 7226 bp in length
 196338 196437: gap of 100 bp
 196438 199835: contig of 3398 bp in length
 199836 199935: gap of 100 bp
 199936 202963: contig of 3028 bp in length
 202964 203063: gap of 100 bp
 203064 206573: contig of 3510 bp in length
 206574 206773: gap of 100 bp
 206774 209337: contig of 2664 bp in length
 209338 209437: gap of 100 bp
 209438 220044: contig of 10607 bp in length
 220045 220144: gap of 100 bp
 220145 224133: contig of 3989 bp in length
 224134 224233: gap of 100 bp
 224234 228476: contig of 2243 bp in length
 228477 226576: gap of 100 bp
 226577 231337: contig of 4761 bp in length
 231338 231437: gap of 100 bp
 231438 234031: contig of 2594 bp in length
 234032 234131: gap of 100 bp
 234132 236913: contig of 2782 bp in length.
 Location/Qualifiers
 1..236913
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="13"
 /clone="RP11-324K20"
 /clone_lib="RPC1-11.2"
 1..4991
 /note="assembly_fragment:00518
 fragment_chain:1"
 5092..7436
 /note="assembly_fragment:00121
 fragment_chain:1"
 7537..25274
 /note="assembly_fragment:00704
 fragment_chain:1"
 25375..28650
 /note="assembly_fragment:00589
 fragment_chain:1"
 28751..31057
 /note="assembly_fragment:01387
 fragment_chain:1"
 31158..36458
 /note="assembly_fragment:00814
 fragment_chain:2"
 36559..41612
 /note="assembly_fragment:00002
 fragment_chain:2"
 41713..44933
 /note="assembly_fragment:00853
 fragment_chain:2"
 45034..47911
 /note="assembly_fragment:01260
 fragment_chain:3"
 48012..50422
 /note="assembly_fragment:00867


```

OY 421 GTCGTCTCCAGATTAAGTGGAGAGCTCTGGGCTTTGGGATTAATAGATGATGATTCAG 480
    |||||
DB 421 GTCGTCTCCAGATTAAGTGGAGAGCTCTGGGCTTTGGGATTAATAGATGATGATTCAG 480
OY 481 ATGGTGACACCAATGAGGCTTACCTTACTGTGGGTCAGAGGAGCGGATGATCTTTTCTGT 540
    |||||
DB 481 ATGGTGACCAATTAATGAATTTCCCTACTGTGGGTCAGAGGAGGATGAAACCAATTTCTTGT 540
OY 541 GAGATACAGCTTATTAAGAGTGGGCTGTGACAGACATTCCTTTTGAACACCTCTCTC 600
    |||||
DB 541 GAGATGATTCCTGTGTAAGAGTGGGCTGTGTAAGACATTCCTGTTTGAAGAGGTATA 600
OY 601 TTTGCTGTGCTGTCTTCAAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
    |||||
DB 601 TTTGCTGTGCTGTCTTCAAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
OY 661 TGCATCTAGAGGAGGCTGTGCTCCGAAATACGCTGTGAGGCTGGAAGAAAGCCCTGGCC 720
    |||||
DB 661 CACATCTTAGAGGAGCTGTGCTCAAAATGCACTGTGAGGCTGGAAGAAAGCCCTGGCC 720
OY 721 ACCTGCTCTCCCAACCTTACAGCTGTGACCTCTCTCTATGAGGAGGAGGAGCTTCTATG 780
    |||||
DB 721 ACCTGCTCTCCCAACCTTACAGCTGTGACCTCTCTCTATGAGGAGGAGGAGCTTCTATG 780
OY 781 CTGAGGCTTAGGAGGCTGTGCTCCGAAATACGCTGTGAGGCTGTGATCTTCTTACACA 840
    |||||
DB 781 CTGAGGCTTAGGAGGCTGTGCTCCGAAATACGCTGTGAGGCTGTGATCTTCTTACACA 840
OY 841 GTCTTACTCTCCATGTGAGACCCCTCTTATTAAGCTTGAAGAAATGAGGAGGATGATGGG 900
    |||||
DB 841 GTCTTACTCTCCATGTGAGACCCCTCTTATTAAGCTTGAAGAAATGAGGAGGATGATGGG 900
OY 901 GCACTGAGGAAGGAGGCTGTGACCCGCTGTGAGGATTTGGACAGCAGCACTGA 948
    |||||
DB 901 GCACTGAGGAAGGAGGCTGTGACCCGCTGTGAGGATTTGGACAGCAGCACTGA 948

RESULT 8
AX448545 948 bp DNA linear PAT 03-JUL-2002
LOCUS Sequence 211 from Patent WO0224726.
DEFINITION AX448545
ACCESSION AX448545
VERSION AX448545.1 GI:21697443
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Veithen, A.
TITLE Olfactory and pheromones g-protein coupled receptors
JOURNAL Patent: WO 0224726-A 211 28-MAR-2002;
Chemcom S.A. (BE)
FEATURES
source location/Qualifiers
1. 948
CDS
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. 948
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37591.1"
/db_xref="GI:21697444"
/translation="METWVQSTYDGFELIGFISHTADLVLFVYMAVFTVALCGNV
LILFLYMDPHLHTPMYFELISLMDLMVCTNVRMANFSLGKSSIFVGCGLQI
GLFVCLVSGEGLIGLMAVDYVAISHPLHPIIMQRCLOITGSSMLFGLDGLIO
MVVWNPYCGLRKVNHPFCMLSLKILACVDTLSLEKVFACCVMLFPEFIIVAS
YAHILCTVLOMHSQAQAKRLATCSSHLTAVTLFGAAMPYIYLRPHRYARPSHDKVAS
IFTVVLPMLPLITISLRNREYVAGLRKGLDRCRISQH"
BASE COUNT 176 a 279 c 228 g 265 t
ORIGIN
Query Match 85.5%; Score 810.4; DB 6; Length 948;
Best Local Similarity 90.9%; Pred. No. 4.2e-205;

```

```

Matches 862; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
OY 1 ATGGGAAGATGGGTGAACACAGCTCTACACAGATGCTTCTCTTGGGATCTTTTC 60
    |||||
DB 1 ATGGAGACGTGGGTGAACACAGCTCTACACAGATGCTTCTCTTGGGATCTTTTC 60
OY 61 CACAGCAGACGTGACCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
    |||||
DB 61 CACAGTACTCTGACCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
OY 121 TGTGGAAATGCTCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
    |||||
DB 121 TGTGGAAATGCTCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
OY 181 TACTTCTCTCTGACACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
    |||||
DB 181 TACTTCTCTCTGACACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
OY 241 AAGATGGACGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
    |||||
DB 241 AAGATGGACGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
OY 301 CAATTTGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
    |||||
DB 301 CAATTTGGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
OY 361 TATGACACTACGCTGAGGCTGTGAGGAGCTGTGAGGAGGCTGTGAGGAGCTGTGAGG 420
    |||||
DB 361 TATGACCGCTATGCTGAGGAGCTGTGAGGAGCTGTGAGGAGCTGTGAGGAGCTGTGAGG 420
OY 421 GTCGTCTCCAGATTAAGTGGAGAGCTCTGGGCTTTGGGATTAATAGATGATGATTCAG 480
    |||||
DB 421 GTCGTCTCCAGATTAAGTGGAGAGCTCTGGGCTTTGGGATTAATAGATGATGATTCAG 480
OY 481 ATGGTGACACCAATGAGGCTTACCTTACTGTGAGCTCAAGAGGCTGTGATCTTTTCTGT 540
    |||||
DB 481 ATGGTGACCAATTAATGAATTTCCCTACTGTGGGTCAGAGGAGGATGAAACCAATTTCTTGT 540
OY 541 GAGGTACAGCTTATTAAGAGCTGGGCTGTGACAGACACTTCTCTTTTGAACACCTCTCT 600
    |||||
DB 541 GAGATGCTATTCCTTGTGAAGCTGGGCTGTGTGAGACACATCTCTTGTGAAGGTATATA 600
OY 601 TTTGCTGTGCTGTCTTCAAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
    |||||
DB 601 TTTGCTGTGCTGTCTTCAAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
OY 661 TGCATCTAGAGGAGGCTGTGCTCCGAAATACGCTGTGAGGCTGTGAGAAAGCCCTGGCC 720
    |||||
DB 661 CACATCTTAGAGGAGCTGTGCTCAAAATGCACTGTGAGGCTGTGAGAAAGCCCTGGCC 720
OY 721 ACCTGCTCTCCCAACCTTACAGCTGTGACCTCTCTCTATGAGGAGGAGGAGCTTCTATG 780
    |||||
DB 721 ACCTGCTCTCCCAACCTTACAGCTGTGACCTCTCTCTATGAGGAGGAGGAGCTTCTATG 780
OY 781 CTGAGGCTTAGGAGGCTGTGCTCCGAAATACGCTGTGAGGCTGTGATCTTCTTACACA 840
    |||||
DB 781 CTGAGGCTTAGGAGGCTGTGCTCCGAAATACGCTGTGAGGCTGTGATCTTCTTACACA 840
OY 841 GTCTTACTCTCCATGTGAGACCCCTCTTATTAAGCTTGAAGAAATGAGGAGGATGATGGG 900
    |||||
DB 841 GTCTTACTCTCCATGTGAGACCCCTCTTATTAAGCTTGAAGAAATGAGGAGGATGATGGG 900
OY 901 GCACTGAGGAAGGAGGCTGTGACCCGCTGTGAGGATTTGGACAGCAGCACTGA 948
    |||||
DB 901 GCACTGAGGAAGGAGGCTGTGACCCGCTGTGAGGATTTGGACAGCAGCACTGA 948

RESULT 9
AB065675 1348 bp DNA linear PRI 23-JUL-2002
LOCUS Homo sapiens gene for seven transmembrane helix receptor, complete
DEFINITION cds, isolate:CBRC7TM_238.
ACCESSION AB065675

```


VERSION AB065675.1 GI:21928623
 KEYWORDS Homo sapiens (isolate:CBRC7M_238) DNA.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S., Tsutsumi,S., Aburatani,H., Asai,K. and Akiyama,Y.
 TITLE Genome-wide discovery and analysis of human seven transmembrane helix receptor genes
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1348)
 AUTHORS Suwa,M.
 JOURNAL Direct Submission
 TITLE Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan (E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/, Tel:81-3-3599-8080, Fax:81-3-3599-8081)
 COMMENT This sequence is a seven transmembrane helix receptor candidate predicted from the whole human genome sequences using our automated system that contains programs of gene finding (genecoder), sequence search, motif-domain assignment and transmembrane helix prediction.
 And the sequence is submitted by the collaborative project between [Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST)] and [Genome Science Division, Research Center for Advanced Science and Technology (RCAST), University of Tokyo].
 FEATURES
 source location/Qualifiers
 1..1348
 /organism="Homo sapiens"
 /isolate="CBRC7M_238"
 /db_xref="taxon:9606"
 /chromosome="5"
 201..1148
 /codon_start=1
 /evidence=not_experimental
 /product="seven transmembrane helix receptor"
 /protein_id="BAC05900.1"
 /db_xref="GI:21928624"
 /translation="METWVQSTTDGFFLLIGTSHSADLVLFVSWAVTVALCGNV LLIPLIYMDPHLHPMTFLSGLDMLVCTNVPMAANFISGRSISFVCGIOL GLPVLCVSEGLILGLIMADRYVAISHPLHYPLIMDQVLOITGSSMAGIIDGLIO MGVVMPNPGYGLKRVNHFCEMLSLKLRVCLSEFKVIFACCVFLLPFSIYAS YAHITGVLMHSAQAKKALATGSSHLTAVTFYGAAMFIYLRPHRYRPSHDKYAS IFYVLTLPMLNPULIYSURNEVGALRKGLDRKRIGSQH"
 CDS
 BASE COUNT 267 a 376 c 310 g 395 t
 ORIGIN
 Query Match 85.5%; Score 810.4; DB 9; Length 1348;
 Best Local Similarity 90.9%; Pred. No. 4.2e-205;
 Matches 862; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Db 441 AAGATGGCAGCCAACTTCCTGTCTGGCAGAGTCCATCTCCTTTGTGGCTGTGGCATA 500
 Qy 301 CAATATGGCTTTTGTCTGTCTGTGGATGATGAGGGGCTTGTGGGATCATGGCT 360
 Db 501 CAATATGGCTTTTGTCTGTCTGTGGATGATGAGGGGCTTGTGGGATCATGGCT 560
 Qy 361 TATAGCAGCTACGAGCGGCTTGGACCCACCTTACATCCATCCATCATCAATCAAG 420
 Db 561 TATAGCAGCTATGAGCGGCTTGGACCCACCTTACATCCATCCATCATCAATCAAG 620
 Qy 421 GTCTGTCTCCAGATTACTGGAGCTCTGGGCTTTGGGATATATGAGATGATTGAG 480
 Db 621 GTCTGTCTCCAGATTACTGGAGCTCTGGGCTTTGGGATATATGAGATGATTGAG 680
 Qy 481 ATGTGGCAGCAGTGGGCTTACCTTACTGTGGGCTCAAGAGACCGGATCACTTTTCTGT 540
 Db 681 ATGTGGGATGATGATTTTCCCTACTGTGGCTTGGAGGAGGATGAAACATTTCTGT 740
 Qy 541 GAGGTAGACGCTTATTTGAAAGCTGGGCTGTGACAGACACTTCCCTTTTGGACACCTCTC 600
 Db 741 GAGATGCTATCTGTGTGAGCTGGGCTGTGTAGACACATCCCTTTTGGAGAGTGATA 800
 Qy 601 TTGCTGTCTGTGTCTGTGATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
 Db 801 TTGCTGTCTGTGTCTGTGATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 860
 Qy 661 TGCATCTAGAGGGCTGTGCTCCGAATACGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
 Db 861 CACATTTAGGAGCTGTGCTCAATATGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 920
 Qy 721 ACCTGCT 780
 Db 921 ACCTGCT 980
 Qy 781 CTGAGGCTTAGGCGCTACCGGGGCTTACGCTATGACAGTGGCTTATCTTTACACA 840
 Db 981 CTGAGGCTTAGGCGCTACCGGGGCTTACGCTATGACAGTGGCTTATCTTTACACC 1040
 Qy 841 GTCTTACTCCATGCTGAGACCCCTTATTTACAGCTTGAAGATGGAGATGATGGG 900
 Db 1041 GTCTTACTCCATGCTGAGACCCCTTATTTACAGCTTGAAGATGGAGATGATGGG 1100
 Qy 901 GCACTGAGGAAGGGGCTGAGCGCTGCAAGATTGGACGACGACACTGA 948
 Db 1101 GCACTGAGGAAGGGGCTGAGCGCTGCAAGATTGGACGACGACGACACTGA 1148
 RESULT 10
 AC008620/c 183455 bp DNA linear HTG 07-MAR-2002
 LOCUS Homo sapiens chromosome 5 clone CTB-1A14, WORKING DRAFT SEQUENCE,
 DEFINITION 8 ordered pieces.
 AC008620
 AC008620.8 GI:19224690
 HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEPIN.
 VERSION Homo sapiens.
 KEYWORDS Homo sapiens
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 183455)
 AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 5
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 183455)
 AUTHORS DOE Joint Genome Institute.
 JOURNAL Direct Submission
 TITLE Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3 (bases 1 to 183455)
 AUTHORS DOE Joint Genome Institute.
 JOURNAL Direct Submission
 TITLE Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Best Local Similarity 91.0%; Pred. No. 9.9e-205;
Matches 860; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```
OY 1 ATGGGAAGATGGGGAACAGCTCTACACAGATGGGCTTCTCCCTGGGCACTTTTCC 60
Db 1 ATGGAGACGTGGGTGAACAGCTCTACACAGATGGGCTTCTCCCTGGGCACTTTTCC 60
OY 61 CACAGCAGACGTAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db 61 CACAGTACTCTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
OY 121 TGTGGGAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 121 TGTGGGAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
OY 181 TACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 181 TACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
OY 241 AGATGGACAGCAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 241 AGATGGACAGCAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
OY 301 CAATTTGGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 301 CAATTTGGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
OY 361 TATGACCACTACGAGCGCTTGAACAGCACTTCTCTCTCTCTCTCTCTCTCT 420
Db 361 TATGACCACTACGAGCGCTTGAACAGCACTTCTCTCTCTCTCTCTCTCTCT 420
OY 421 GTCGTCTCAGATTTACTGAGAGCTCTGAGGCTTTGGGTAATAGATGAGATT 480
Db 421 GTCGTCTCAGATTTACTGAGAGCTCTGAGGCTTTGGGTAATAGATGAGATT 480
OY 481 ATGTGGACAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 481 ATGTGGACAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
OY 541 GAGGTACAGCTTATTTAGAGCTGAGCTGAGAGCACTTCTCTCTCTCTCTCT 600
Db 541 GAGGTACAGCTTATTTAGAGCTGAGCTGAGAGCACTTCTCTCTCTCTCTCT 600
OY 601 TTTGCTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Db 601 TTTGCTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
OY 661 TGCATCTAGAGGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Db 661 TGCATCTAGAGGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
OY 721 ACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db 721 ACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
OY 781 CTGAGGCTCTAGGCGCTACCGGGCCCTTAGCCATGACAAGGTGAGCTCTAT 840
Db 781 CTGAGGCTCTAGGCGCTACCGGGCCCTTAGCCATGACAAGGTGAGCTCTAT 840
OY 841 GTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 841 GTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
OY 901 GCAGTGAAGAGGCGCTGAGCGCTGAGAGATGAGGAGCCAGCAGC 945
Db 901 GCAGTGAAGAGGCGCTGAGCGCTGAGAGATGAGGAGCCAGCAGC 945
```

RESULT 12
LOCUS AX241793 945 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 541 from Patent W00127158.
ACCESSION AX241793

VERSION AX241793.1 GI:15798668
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 945)
AUTHORS Bellenson,J., Smith,D., Lancel,D., Glusman,G., Fuchs,T. and Yanai,I.
TITLE Olfactory receptor sequences
JOURNAL Patent: WO 0127158-A 541 19-APR-2001;
FEATURES
source location/Qualifiers
1..945
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="(H389390 nucleotide)"
BASE COUNT 175 a 279 c 227 g 264 t
ORIGIN

Query Match 85.2%; Score 807.4; DB 6; Length 945;
Best Local Similarity 90.9%; Pred. No. 2.6e-204;
Matches 859; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

```
OY 1 ATGGGAAGATGGGTAACACAGCTCTACACAGATGGCTTCTCTCTGGCACTTTTCC 60
Db 1 ATGGAGACGTGGGTGAACAGCTCTACACAGATGGCTTCTCTCTGGCACTTTTCC 60
OY 61 CACAGCAGACGTAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db 61 CACAGTACTCTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
OY 121 TGTGGGAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 121 TGTGGGAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
OY 181 TACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 181 TACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
OY 241 AGATGGACAGCAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 241 AGATGGACAGCAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
OY 301 CAATTTGGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 301 CAATTTGGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
OY 361 TATGACCACTACGAGCGCTTGAACAGCACTTCTCTCTCTCTCTCTCTCTCT 420
Db 361 TATGACCACTACGAGCGCTTGAACAGCACTTCTCTCTCTCTCTCTCTCTCT 420
OY 421 GTCGTCTCAGATTTACTGAGAGCTCTGAGGCTTTGGGTAATAGATGAGATT 480
Db 421 GTCGTCTCAGATTTACTGAGAGCTCTGAGGCTTTGGGTAATAGATGAGATT 480
OY 481 ATGTGGACAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 481 ATGTGGACAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
OY 541 GAGGTACAGCTTATTTAGAGCTGAGCTGAGAGCACTTCTCTCTCTCTCTCT 600
Db 541 GAGGTACAGCTTATTTAGAGCTGAGCTGAGAGCACTTCTCTCTCTCTCTCT 600
OY 601 TTTGCTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Db 601 TTTGCTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
OY 661 TGCATCTAGAGGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Db 661 TGCATCTAGAGGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
OY 721 ACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db 721 ACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
```

```

Db      721 ACCTGCTCTCCACCTGACAGCTGCACCTCTTCTATGGGAGCAGCATTCATCTAC 780
Oy      781 CTGAGGCTTACGCTACCGGCGCCCTAGCAGACAAAGTGCGCTCTATCTTACACA 840
Db      781 CTGAGGCTTACGCTACCGGCGCCCTAGCAGACAAAGTGCGCTCTATCTTACACA 840
Oy      841 GTCTTACTCTCCATGCTGACACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGG 900
Db      841 GTCTTACTCTCCATGCTGACACCCCTCATTTACAGCTTGAGGAATGGGAGGTGATGGG 900
Oy      901 GCACCTGAGGAGGCGTGGACCGCTGAGGATGGCGAGCAGCAGC 945
Db      901 GCACCTGAGGAGGCGTGGACCGCTGAGGATGGCGAGCAGCAGC 945

RESULT 13
AC103024 106325 bp DNA linear HTG 13-JUL-2002
LOCUS Rattus norvegicus clone CH230-199E4, *** SEQUENCING IN PROGRESS
DEFINITION *** 58 unordered pieces.
AC103024 AC103024.3 GI:21730133
VERSION HTG: HTGS PHASE1.
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 106325)
Muzny,D.M., Adams,C., Adio-oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,C.,
Barberia,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouch,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Doutwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamthair,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs,J.F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsone,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherrer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,M.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umanai,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleceyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Morley,K.C.

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT
Direct Submission
Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 106325)
Morley,K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:17974430.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GIW
Center clone name: CH230-199E4
----- Summary Statistics
Sequencing vector: Plasmid.
Chemistry: Dye-terminator Big Dye. 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 54989 bases at least Q40
Consensus quality: 61953 bases at least Q30
Consensus quality: 67827 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_delta.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 58 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1013: contig of 1013 bp in length
1114: gap of unknown length
1114: contig of 1132 bp in length
2246: gap of unknown length
2345: gap of 1329 bp in length
3674: contig of 1329 bp in length
3675: gap of unknown length
3775: contig of 1544 bp in length
5318: gap of unknown length
5419: contig of 1375 bp in length
6794: gap of unknown length
6893: gap of 1464 bp in length
8357: contig of 1464 bp in length
8358: gap of unknown length
8457: gap of 1221 bp in length
9678: contig of 1142 bp in length
9679: gap of unknown length
10920: contig of 1142 bp in length
10921: gap of unknown length
11021: gap of 1466 bp in length
11021: contig of 1466 bp in length
12486: gap of unknown length
12487: gap of 1144 bp in length
13730: contig of 1144 bp in length
13731: gap of unknown length
15022: contig of 1192 bp in length
15023: gap of unknown length
15122: gap of unknown length
15123: contig of 1112 bp in length
16234: gap of unknown length
16335: gap of 1538 bp in length
17872: contig of 1538 bp in length
17873: gap of unknown length
17973: contig of 1398 bp in length
19371: gap of unknown length
19371: contig of 1237 bp in length
20707: gap of unknown length
20807: gap of 2169 bp in length
22976: contig of 2169 bp in length
23075: gap of unknown length
23077: contig of 1198 bp in length
24374: gap of unknown length
24375: contig of 1140 bp in length
25514: contig of 1140 bp in length
25515: gap of unknown length

```



```

VERSION      AY073214.1  GI:18479725
KEYWORDS     Mus musculus.
SOURCE       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ORGANISM     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE    1 (bases 1 to 948)
AUTHORS      Zhang, X. and Firestein, S.
TITLE        The olfactory receptor gene superfamily of the mouse
JOURNAL      Nat. Neurosci. 5 (2), 124-133 (2002)
MEDLINE      21678663
PUBMED       11802173

REFERENCE    2 (bases 1 to 948)
AUTHORS      Young, J. M., Friedman, C., Williams, E. M., Ross, J. A., Tonnes-Priddy, L.
TITLE        and Trask, B. J.
JOURNAL      Different evolutionary processes shaped the mouse and human
MEDLINE      olfactory receptor gene families
PUBMED       Hum. Mol. Genet. 11 (5), 535-546 (2002)
PUBMED       21864068

REFERENCE    3 (bases 1 to 948)
AUTHORS      Adams, M.
TITLE        Submitted
JOURNAL      Direct Submission
MEDLINE      Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,
PUBMED       Rockville, MD 20850, USA

FEATURES     Location/Qualifiers
             1..948
             /organism="Mus musculus"
             /db_xref="taxon:10090"
             /chromosome="11"
             /map="11B1"
             /note="GA_X5JB7W6C8B-402233-401286"
             /product="Olfactory receptor MOR276-1"
             1..948
             /codon_start=1
             /evidence="not experimental"
             /product="Olfactory receptor MOR276-1"
             /protein_id="AAL60877.1"
             /translation="MGIMLNSSVDGFIILGIFSOSQDILLSTVMLVETVALGNAV
             LLILLYTDRLRHPYFELSQLSLMDLVNCPKMAVNFSGRSISFAGCGIOI
             GFEPYVSGSLGILMAVDYRYAISHPLPIIMSGKVLQINGSSMAGFIDGIIIO
             MYAMSLPYGSGSRIDHFCEVPAALLKACADPSLPTLLFACQVEMLIPSTIVS
             YARLIGAVLRHNSHQRKALATCSHLTVSLFYGAAMFTYLRPRRYRAPSHDKVVS
             IFYTVLRPMNPLIYSLRNREVMGALKRGDRCRVGSOH"

BASE COUNT   182 a      275 c      213 g      278 t

Query Match      78.4%; Score 743.2; DB 10; Length 948;
Best Local Similarity 86.5%; Pred. No. 3.4e-187;
Matches 820; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

```

```

QY 301 CAATGGCGTTTGTGTCCTGTTGGGATCTGAGGGGCTTGTGGGACTATGGCT 360
DB 301 CAATGGGATTTTGTGTCCTGTTGGGATCTGAGGGGCTTGTGGGACTATGGCT 360
QY 361 TATGACACTACGTGCGCGTTAGCCACCCACTTACTATCCATCTCATGAAATCAGAG 420
DB 361 TATGATGCTATGCGCATATAGCACCCACTTACTATCCATCTCATGAAATCAGAG 420
QY 421 GTCTGTCTCCAGATTTATGAGGAGCTCTGGGCGCTTGGGTAATATGATGAGATCAG 480
DB 421 GTCTGTCTCCAGATTTATGAGGAGCTCTGGGCGCTTGGGTAATATGATGAGATCAG 480
QY 481 ATGGTGACCCAGATGCGGTACTTACTGTGGGCTCAAGAGGCGTGATCACTTTTCTGT 540
DB 481 ATGGTGACCCAGATGCGGTACTTACTGTGGGCTCAAGAGGCGTGATCACTTTTCTGT 540
QY 541 GAGGTACAGCTTTATTTGAAGCTGCGCTGTGACAGACCTTCCCTTTTGTACACCTCTC 600
DB 541 GAGGTACAGCTTTATTTGAAGCTGCGCTGTGACAGACCTTCCCTTTTGTACACCTCTC 600
QY 601 TTTGCTGTCTGTCTTATGATGCTCTTCTCCCTTCATCATGATGATGATGATGATGAT 660
DB 601 TTTGCTGTCTGTCTTATGATGCTCTTCTCCCTTCATCATGATGATGATGATGATGAT 660
QY 661 TGCATCTAGGGGCTGTGCTCCGAATACGCTCTGCTCAGGCTGGAAGAAAGCCCTGGCC 720
DB 661 CGCATCTTGGGGGCTGTGCTCCGAATACGCTCTGCTCAGGCTGGAAGAAAGCCCTGGCC 720
QY 721 ACCTGCTCTCCACCTACACGCTGTACACCTCTTCTATGAGGAGGAGGAGGAGGAGGAG 780
DB 721 ACTTGTCTCTCCACCTACACGCTGTACACCTCTTCTATGAGGAGGAGGAGGAGGAGGAG 780
QY 781 CTGAGGCGCTAGGCGCTACCGGGGCGCTAGGCGATGACAAGGAGGAGGAGGAGGAGGAG 840
DB 781 CTGAGGCGCAAGGCGATGCGGCTCTTACGATGACAAGGAGGAGGAGGAGGAGGAGGAG 840
QY 841 GTCTTACTCCATGCTGACACCCCTCATTTTACAGCTTGAAGAAATGGGAGGATGAGGG 900
DB 841 GTTCTTACTCTTATGCTGACACCCCTCATTTTATAGCTTGAAGAAAGGAGGATGAGGG 900
QY 901 GCACTGAGGAAGGGGCTGAGACCGCTGCAAGGATTTGGCAGCCAGCACTGA 948
DB 901 GCACTGAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 948

```

```

RESULT 15
AL645688      123115 bp      DNA      linear      ROD 05-APR-2002
LOCUS        Mouse DNA sequence from clone Rp23-54N20 on chromosome 11, complete
DEFINITION   sequence.
ACCESSION    AL645688
VERSION      AL645688.12  GI:19715734
KEYWORDS     HTG.
ORGANISM     Mus musculus.
SOURCE       house mouse.
REFERENCE    1
AUTHORS      Howden, P.
TITLE        Direct Submission
JOURNAL      Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquer@anger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Mar 25, 2002 this sequence version replaced gi:1969540.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=

```

30), an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em', EMBL; Sw', SWISSPROT; Tr', TREMBL; Wp', WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep. RP23-54N20 is from the RPCI-23 Mouse PNC library constructed by the group of Pieret de Jong. For further details see <http://www.chori.org/bacpac/home.htm>. VECTOR: pBACE3.6.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .123115 |

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 20:23:29 ; Search time 267 Seconds
(without alignments)
7995.859 Million cell updates/sec

Title: US-09-634-109D-1

Perfect score: 948
Sequence: 1 atggagaagtgggtgaccca.....ggatggcagccagcactga 948

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-----------------------------|
| 1 | 948 | 100.0 | 948 | 22 | AAS00555 Human CON167 G pro |
| 2 | 948 | 100.0 | 1070 | 22 | AAS00560 Human CON167 G pro |
| 3 | 940 | 99.2 | 951 | 24 | ABK97214 Human G-protein co |
| 4 | 935.4 | 98.7 | 1120 | 24 | ABO8046 Human GPCR CDNA # |
| 5 | 908.8 | 95.9 | 975 | 22 | AAH31969 Human olfactory r |
| 6 | 810.4 | 85.5 | 948 | 22 | AAS42296 Human CDNA encodin |
| 7 | 810.4 | 85.5 | 948 | 24 | ABK68506 Human DNA for olfa |
| 8 | 810.4 | 85.5 | 948 | 24 | AA29683 Human G-protein co |
| 9 | 810.4 | 85.5 | 948 | 24 | ABK37582 DNA encoding G-cou |

| | | | | | |
|----|-------|------|------|----|-----------------------------|
| 10 | 809 | 85.3 | 945 | 22 | AAH31875 Human olfactory re |
| 11 | 807.4 | 85.2 | 945 | 22 | AAH31968 Human olfactory re |
| 12 | 672.4 | 70.9 | 965 | 24 | AA26374 Human G-protein co |
| 13 | 526.6 | 55.5 | 669 | 21 | AAH51179 Murine olfactory r |
| 14 | 526.6 | 55.5 | 669 | 21 | AAH51183 Murine olfactory r |
| 15 | 525 | 55.4 | 669 | 21 | AAH51181 Murine olfactory r |
| 16 | 390.2 | 41.2 | 488 | 22 | AAH84008 Gorilla olfactory |
| 17 | 389.4 | 41.1 | 939 | 24 | ABK68668 Human DNA for olfa |
| 18 | 388.8 | 41.0 | 1290 | 19 | AAH53208 Human olfactory ol |
| 19 | 386.2 | 40.7 | 936 | 24 | ABK68667 Human DNA for olfa |
| 20 | 379.8 | 40.1 | 939 | 24 | ABK68670 Human DNA for olfa |
| 21 | 371 | 39.1 | 963 | 24 | ABK68671 Human DNA for olfa |
| 22 | 366.2 | 38.6 | 963 | 24 | ABK68669 Human DNA for olfa |
| 23 | 362.2 | 38.2 | 974 | 24 | ABH89126 Human GPCR7c nucle |
| 24 | 361.8 | 38.2 | 975 | 22 | AAS42423 Human CDNA encodin |
| 25 | 361.8 | 38.2 | 975 | 24 | ABK68567 Human DNA for olfa |
| 26 | 361.8 | 38.2 | 975 | 24 | AA29682 Human G-protein co |
| 27 | 361.8 | 38.2 | 975 | 24 | ABK37709 Human GPCR7a nucle |
| 28 | 360.6 | 38.0 | 951 | 22 | AAS42447 Human CDNA encodin |
| 29 | 360.6 | 38.0 | 951 | 24 | ABH89124 Human GPCR7a nucle |
| 30 | 360.6 | 38.0 | 951 | 24 | ABK37733 Human GPCR7a nucle |
| 31 | 352.4 | 37.2 | 964 | 22 | AAH31643 Human olfactory re |
| 32 | 351.8 | 37.1 | 982 | 24 | ABH89125 Human GPCR7b nucle |
| 33 | 349.6 | 36.9 | 870 | 24 | ABK68556 Human DNA for olfa |
| 34 | 346.6 | 36.6 | 948 | 24 | ABH89121 Human GPCR6a nucle |
| 35 | 345 | 36.4 | 945 | 22 | AAH31912 Human olfactory re |
| 36 | 345 | 36.4 | 948 | 22 | AAS42426 Human CDNA encodin |
| 37 | 345 | 36.4 | 948 | 24 | ABH89123 Human GPCR6c nucle |
| 38 | 345 | 36.4 | 948 | 24 | ABK68568 Human DNA for olfa |
| 39 | 345 | 36.4 | 948 | 24 | ABK37712 Human GPCR7a nucle |
| 40 | 345 | 36.4 | 949 | 24 | ABH89122 Human GPCR6a nucle |
| 41 | 344.8 | 36.4 | 963 | 24 | ABO88381 Human G-protein co |
| 42 | 343.8 | 36.3 | 990 | 22 | AAH31892 Human olfactory re |
| 43 | 343.8 | 36.3 | 993 | 24 | ABK68569 Human DNA for olfa |
| 44 | 343.8 | 36.3 | 948 | 22 | AAS42425 Human CDNA encodin |
| 45 | 343.8 | 36.3 | 1047 | 22 | ABK37711 DNA encoding G-cou |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| AAS00555 | AAS00555 standard; DNA: 948 BP. |
| XX | XX |
| AC | AAS00555; |
| XX | XX |
| DT | 14-MAY-2001 (first entry) |
| XX | XX |
| DE | Human CON167 G protein seven transmembrane receptor DNA #1. |
| XX | XX |
| KW | CON167; G protein coupled receptor; transmembrane receptor; ds; |
| KW | neurological disorder; psychiatric disease; schizophrenia; depression; |
| KW | anxiety; bipolar disorder; affective disorder; epilepsy; neuritis; |
| KW | attention deficit hyperactivity disorder; ADHD; neurotransmitter; neuroptahy; |
| KW | neurosis; Alzheimer's disease; Parkinson's disease; migraine; therapy; |
| KW | senile dementia; hybridisation assay; diagnosis; transgenic animal. |
| XX | XX |
| OS | Homo sapiens. |
| XX | XX |
| FH | Key |
| FT | Location/Qualifiers |
| FT | 1..948 |
| FT | /*tag- a |
| FT | /product= "CON167" |
| XX | XX |
| PN | WO200114554-A1. |
| XX | XX |
| PD | 01-MAR-2001. |
| XX | XX |
| PF | 08-AUG-2000; 2000WO-US21566. |
| XX | XX |
| PR | 19-AUG-1999; 99US-0377563. |
| XX | XX |

PT New G-protein coupled receptor polypeptides, encoding nucleic acids and
 PT immunospecific antibodies, useful for treating cardiomyopathy,
 PT atherosclerosis, diabetes or disorder related to cell signal processing
 PT and metabolic pathway modulation -
 XX
 PS Claim 9; Page 86; 146pp; English.

XX The invention relates to a novel nucleic acid molecule comprising a
 CC nucleic acid sequence encoding a G-protein coupled receptor (GPCR)
 CC polypeptide. The proteins of the invention have cardiant,
 CC antiatherosclerotic, antidiabetic, hypotensive, anorectic, cyostatic,
 CC immunosuppressive, neuroprotective, nootropic, antiparkinsonian, and
 CC antileptic activity. The nucleic acids of the invention may have a use
 CC in gene therapy. The sequences of the invention are useful for treating
 CC or preventing a GPCR-associated disorder, preferably cardiomyopathy,
 CC atherosclerosis or a disorder related to cell signal processing and
 CC metabolic pathway modulation. An antibody of the invention is useful for
 CC treating diabetes or a disorder related to cell signal processing and
 CC metabolic pathway modulation. GPCR-associated disorders include
 CC hypertension, obesity, cancer, graft versus host disease, Alzheimer's
 CC disease, Parkinson's disorder and dyslipidaemia. The sequences shown in
 CC AB088043-AB088093 encode the GPCR proteins of the invention, shown in
 CC ABP01129-ABP61179.

XX Sequence 1120 BP; 210 A; 318 C; 268 G; 324 T; 0 other;

Query Match 98.7%; Score 935.4; DB 24; Length 1120;

Best Local Similarity 99.8%; Pred. No. 1.9e-252; Mismatches 1; Indels 1; Gaps 1;

Matches 947; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AAGGAGATGGGTGAACACCTCTACACAGATGGCTTCTTCCCTTGGGACATCTTCC 60
 DB 3 AAGGAGATGGGTGAACACCTCTACACAGATGGCTTCTTCCCTTGGGACATCTTCC 62
 QY 61 CACAGCCACACGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
 DB 63 CACAGCCACACGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122
 QY 121 TGTGGGAATGTCCT 180
 DB 123 TGTGGGAATGTCCT 182
 QY 181 TACTTCTTCTCAGCAGCT 240
 DB 183 TACTTCTTCTCAGCAGCT 242
 QY 241 AAGATGGCAGCCACTTCT 300
 DB 243 AAGATGGCAGCCACTTCT 302
 QY 301 CAAATTTGGCTTTTGTCT 360
 DB 303 CAAATTTGGCTTTTGTCT 362
 QY 361 TATGACCACTAGTGGCGCTTACACACCACTTCACTATCCATCCCTCTCTCTCTCTCT 420
 DB 363 TATGACCACTAGTGGCGCTTACACACCACTTCACTATCCATCCCTCTCTCTCTCTCT 422
 QY 421 GTCTGTCTCTCAATTAATGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
 DB 423 GTCTGTCTCTCAATTAATGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
 QY 481 ATGTGGGCGAGGCTTACCTTACTGTGCTCAAGGAGGTGGATACATTTTCTCTCT 540
 DB 483 ATGTGGGCGAGGCTTACCTTACTGTGCTCAAGGAGGTGGATACATTTTCTCTCTCT 542
 QY 541 GAGGTACAGCTTTATTTAGAGCTGGCGCTTGACACACTTCCCTTTTGGACACCTCTCT 600
 DB 543 GAGGTACAGCTTTATTTAGAGCTGGCGCTTGACACACTTCCCTTTTGGACACCTCTCT 602
 QY 601 TTTGGCTGTGTCT 660
 DB 603 TTTGGCTGTGTCT 662

QY 661 TCCATC-CTAGGCGCTGTGCTCCGAATACGCTGTGCTAGGCGCTGAAAAAAGCCCTGGC 719
 DB 663 TCCATCTAGGCGGCTGTGCTCCGAATACGCTGTGCTAGGCGCTGAAAAAAGCCCTGGC 722
 QY 720 CACCTGCTCTCTCCACCTACACAGCTGTACCTCTTTATTTGGGCGACCATCTTATGTA 779
 DB 723 CACCTGCTCTCTCCACCTACACAGCTGTACCTCTTTATTTGGGCGACCATCTTATGTA 782
 QY 780 CCGTGGGCGCTTACCGGCGCTTACCGCATGACAGGCGCTTATTTTATAC 839
 DB 783 CCGTGGGCGCTTACCGGCGCTTACCGCATGACAGGCGCTTATTTTATAC 842
 QY 840 ACTCTTACTCTCCATGCTGACCCCTCTATTTACAGCTTGAAGATGGGAGCTGATGG 899
 DB 843 ACTCTTACTCTCCATGCTGACCCCTCTATTTACAGCTTGAAGATGGGAGCTGATGG 902
 QY 900 GGCACCTGAGGAAGGCGCTGACCGCTGACAGATTTGGCAGCCAGCACTGA 948
 DB 903 GGCACCTGAGGAAGGCGCTGACCGCTGACAGATTTGGCAGCACTGA 951

RESULT 5

AAH31969 standard; DNA; 975 BP.

AAH31969;

30-JUL-2001 (first entry)

Human olfactory receptor polynucleotide, SEQ ID NO: 542.

Human: olfactory receptor; OR: primary scent determination;
 secondary scent determination; polypeptide library; odour receptor;
 scent profile; scent fingerprint; scent representation; ds.

Homosapiens.

W0200127158-A2.

19-APR-2001.

06-OCT-2000; 2000WO-US27582.

08-OCT-1999; 99US-0158615.

24-FEB-2000; 2000US-0184809.

(DIGI-) DIGISCENTS.

(YEDA) YEDA RES & DEV CO LTD.

Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

WPI; 2001-290713/30.

New polynucleotides which encode polypeptides involved in olfactory
 sensation for identifying olfactory agonists and antagonists -

Claim 8; Page 397; 1857pp; English.

The present sequence is one of a number of isolated polynucleotides
 which encode polypeptides involved in olfactory sensation. The
 polynucleotides can be used in screening for olfactory agonists and
 antagonists. The methods allow for the determination of primary
 scents and the identification of the odour receptors used to detect
 these primary scents. The methods also enable determination of
 secondary scents and the identification of combinations of odour
 receptors that are involved in detecting such secondary scents.
 This enables the construction of a scent representation (also called
 a scent fingerprint or scent profile), which may be used to re-create
 and edit scents. Libraries of olfactory receptors are useful for
 determining the interaction pattern of a composition with the receptors,
 and can be used for determining differences in the olfactory faculties
 of different individuals.

QY 181 TACTTCTCTCAGACAGCTCTCCCTCATGAGACCTATGTTGCTGTGTAACATTGTGCA 240
 Db 181 TACTTCTCTCAGACAGCTCTCCCTCATGAGACCTATGTTGCTGTGTAACATTGTGCA 240
 QY 241 AAGATGAGGAGCAACTTCTCTGTGAGAGAAATCCATCTCTTGTGGGCTGTGGCATA 300
 Db 241 AAGATGAGGAGCAACTTCTCTGTGAGAGAAATCCATCTCTTGTGGGCTGTGGCATA 300
 QY 301 CAATATGGCTTTTGTGTCTGTGTGGGATCTGAGGGGCTTGTGTGGGACTATGGCT 360
 Db 301 CAATATGGCTTTTGTGTCTGTGTGGGATCTGAGGGGCTTGTGTGGGACTATGGCT 360
 QY 361 TATGACACACTAGTGGCCGTAGGACACCACTTATCCATCCATCCATCAATCAAGAG 420
 Db 361 TATGACACACTAGTGGCCGTAGGACACCACTTATCCATCCATCCATCAATCAAGAG 420
 QY 421 GTCTGTCTCCAGATTACTGAGAGCTCTGAGCCCTTTGGGATTAATGATGAGTATTCAG 480
 Db 421 GTCTGTCTCCAGATTACTGAGAGCTCTGAGCCCTTTGGGATTAATGATGAGTATTCAG 480
 QY 481 ATGGGAGGAGCATGAGGCTTACTGCTGATGAGGCTCAAGAGGAGGAGTATCTTTTCTGT 540
 Db 481 ATGGGAGGAGCATGAGGCTTACTGCTGATGAGGCTCAAGAGGAGGAGTATCTTTTCTGT 540
 QY 541 GAGGTACAGCTTTATTTGAAAGCTGAGCTGTGACAGACATCTCTTTTGTGACACCTCTC 600
 Db 541 GAGGTACAGCTTTATTTGAAAGCTGAGCTGTGACAGACATCTCTTTTGTGACACCTCTC 600
 QY 601 TTTGCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 660
 Db 601 TTTGCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 660
 QY 661 TGCATCCAGAGGAGGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 720
 Db 661 TGCATCCAGAGGAGGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 720
 QY 721 ACCTGCTCTCTCCACCTTACAGCTGTCAACCTCTTCTTATGAGGAGGAGGAGGAGGAGG 780
 Db 721 ACCTGCTCTCTCCACCTTACAGCTGTCAACCTCTTCTTATGAGGAGGAGGAGGAGGAGG 780
 QY 781 CTGAGGCTTACGAGCTGTCAACCTCTTCTTATGAGGAGGAGGAGGAGGAGGAGGAGG 840
 Db 781 CTGAGGCTTACGAGCTGTCAACCTCTTCTTATGAGGAGGAGGAGGAGGAGGAGGAGG 840
 QY 841 GTCTTACTTCCATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
 Db 841 GTCTTACTTCCATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
 QY 901 GCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 948
 Db 901 GCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 948
 RESULT 7
 ABK68506
 ID ABK68506 standard; DNA; 948 BP.
 XX ABK68506:
 AC 02-JUL-2002 (first entry)
 DT Human DNA for olfactory and pheromone G protein-coupled receptor #106.
 DE Human: ds; gene: olfactory and pheromone G protein coupled receptor;
 KW GPCR; tranquilizer; antidepressant; neuroleptic; endocrine; anabolic;
 KW anorectic; taste; fragrance; food additive; cosmetic; cell migration;
 KW sterility; psychotic disorder; neurological disorder; anxiety;
 KW schizophrenia; manic depression; depression; axonal growth;
 KW menstrual cycle; appetite sexual motivation; sexual attraction;
 KW aggression.
 XX Homo sapiens.
 OS

XX WO200224726-A2.
 PN 28-MAR-2002.
 PD 21-SEP-2001; 2001MO-BE00162.
 PF 22-SEP-2000; 2000EP-0870211.
 PR (CHEM-) CHEMCOM SA.
 PA Veithen A;
 PI WPI: 2002-330013/36.
 DR P-PDSB; AAU95619.
 RX Novel pheromone G-protein coupled receptor and receptor-derived
 PT agonists, antagonists or inhibitors useful in food or cosmetic products
 PT or in the treatment or prevention of neurological disorders such as
 PT anxiety and schizophrenia -
 PS Disclosure; Page 330-331; 833pp; English.
 XX The invention relates to olfactory and pheromone G-protein coupled
 CC receptor (GPCR) or a protein 95% identical to the GPCR, a specific active
 CC portion and its encoding polynucleotide. Also included are an agonist,
 CC antagonist or inhibitor of the GPCR or the polynucleotide, a vector
 CC comprising the polynucleotide, a cell transformed by the vector, a non-
 CC human mammal comprising a partial or total deletion of the polynucleotide
 CC encoding the receptor and screening (detection and possibly, recovering)
 CC of compounds which are known or not known to be agonist, antagonists or
 CC inhibitors of natural compounds to the GPCR. The receptor-derived
 CC agonists, antagonists, inhibitors or compounds are used as an
 CC improvement, elimination or substitution of an existing taste and/or a
 CC fragrance of (or in) the food and/or cosmetic products. They can also be
 CC used in the preparation of medicament in the treatment and/or prevention
 CC of a mammalian disorder, such as cell migration, sterility, psychotic and
 CC neurological disorders, including anxiety, schizophrenia, manic
 CC depression, depression, for promoting axonal growth, nerve cell
 CC connection and nerve regeneration for modulating male and female
 CC endocrine functions, hormone production and the menstrual cycle, for the
 CC prevention or the treatment by stimulation of several mammalian
 CC behaviours, such as stimulation or suppression of appetite, sexual
 CC motivation, sexual attraction, aggression and for promoting or
 CC suppressing chemical communication between organisms. The present
 CC sequence is a human DNA encoding an olfactory and pheromone GPCR.
 CC XX
 SQ Sequence 948 BP; 176 A; 279 C; 228 G; 265 T; 0 other:
 Query Match 85.5%; Score 810.4; DB 24; Length 948;
 Best Local Similarity 90.9%; Pred. No. 2.3e-217;
 Matches 862; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
 QY 1 ATGGGAGATGGGTGACCACTCTTACACAGATGGCTTCTCTTGGGAGCTTTTTC 60
 Db 1 ATGGGAGATGGGTGACCACTCTTACACAGATGGCTTCTCTTGGGAGCTTTTTC 60
 QY 61 CACAGCAGACTGACCTGTCTCTCTCTGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
 Db 61 CACAGTACTGTGACCTGTCTCTCTCTGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
 QY 121 TGTGGGATGTCCTCTCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
 Db 121 TGTGGGATGTCCTCTCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
 QY 181 TACTTCTCTCAGACAGCTTCTCCCTCATGAGACCTATGTTGCTGTGTAACATTGTGCA 240
 Db 181 TACTTCTCTCAGACAGCTTCTCCCTCATGAGACCTATGTTGCTGTGTAACATTGTGCA 240
 QY 241 AAGATGAGGAGCAACTTCTCTGTGAGAGAAATCCATCTCTTGTGGGCTGTGGCATA 300
 Db 241 AAGATGAGGAGCAACTTCTCTGTGAGAGAAATCCATCTCTTGTGGGCTGTGGCATA 300

QY 301 CAATTGGCTTTTCTCTCTCTGAGATCTGAGGGCTCTTGGTGGACATGAGCT 360
 DB 301 CAATTGGCTTTTCTCTCTCTGAGATCTGAGGGCTCTTGGTGGACATGAGCT 360
 QY 361 TATGACCACTAGTGGCGGCTTATGACCACTGACATGACATGACATGACATGAC 420
 DB 361 TATGACCACTAGTGGCGGCTTATGACCACTGACATGACATGACATGACATGAC 420
 QY 421 GTCCTGCTCAGATTAAGTGGAGCTTGGGCTTGGGATATGATGATGATGATGAT 480
 DB 421 GTCCTGCTCAGATTAAGTGGAGCTTGGGCTTGGGATATGATGATGATGATGAT 480
 QY 481 ATGCTGGGACCACTGGGCTTACTTACTGCTGCAAGAGCGTGGATCACTTTTCTGT 540
 DB 481 ATGCTGGGACCACTGGGCTTACTTACTGCTGCAAGAGCGTGGATCACTTTTCTGT 540
 QY 541 GAGGTCAAGACTTATTAAGTGGAGCTGACACACTTCCCTTTTGGACACCTTCCTC 600
 DB 541 GAGGTCAAGACTTATTAAGTGGAGCTGACACACTTCCCTTTTGGACACCTTCCTC 600
 QY 601 TTTGCTGCTGCTCTTCAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
 DB 601 TTTGCTGCTGCTCTTCAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
 QY 661 TGCATCTGAGGGCTGCTGCTGCAATAGCTGCTGCAAGAGCGTGGATCACTTTTCTGT 720
 DB 661 TGCATCTGAGGGCTGCTGCTGCAATAGCTGCTGCAAGAGCGTGGATCACTTTTCTGT 720
 QY 721 ACCTGCT 780
 DB 721 ACCTGCT 780
 QY 781 CTGAGGCTTACGAGGCTGAGGCGGCTTACGAGGCGGCTTACGAGGCGGCTTAC 840
 DB 781 CTGAGGCTTACGAGGCTGAGGCGGCTTACGAGGCGGCTTACGAGGCGGCTTAC 840
 QY 841 GTCCTGCTCAGATTAAGTGGAGCTTGGGCTTGGGATATGATGATGATGATGAT 900
 DB 841 GTCCTGCTCAGATTAAGTGGAGCTTGGGCTTGGGATATGATGATGATGATGAT 900
 QY 901 GCACGTGAGGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCT 948
 DB 901 GCACGTGAGGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCT 948
 RESULT 8
 AAD29683
 ID AAD29683 standard: cDNA: 948 BP.
 AC AAD29683;
 XX
 DT 17-MAY-2002 (first entry)
 DE Human G-protein coupled receptor (GCRRC-17) cDNA.
 XX
 KM Human: G-protein coupled receptor: GCRRC-17: cell proliferative disorder;
 KM neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;
 KM metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;
 KM Pick's disease; Huntington's disease; Parkinson's disease; hypertension;
 KM atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;
 KM osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
 KM anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;
 KM transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;
 KM nootropic; neuroprotective; cardiant; immunosuppressive; anorectic;
 KM virulence; gene; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..948
 FT /tag= a
 FT /product= "Human GCRRC-17"
 FT sig_peptide 1..72

FT mat_peptide 73..945
 FT /tag= b
 FT /tag= c
 FT /product= "Human mature GCRRC-17"
 PN MO200210387-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 25-JUL-2001: 2001WO-US23433.
 XX
 PR 27-JUL-2000: 2000US-221478P.
 PR 03-AUG-2000: 2000US-223268P.
 PR 21-AUG-2000: 2000US-227054P.
 PR 08-SEP-2000: 2000US-231121P.
 PR 13-SEP-2000: 2000US-232243P.
 PR 15-SEP-2000: 2000US-232691P.
 PR 22-SEP-2000: 2000US-235146P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR:
 PI Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Walla NK:
 PI Hafalia Afa, Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L:
 PI Graul RC, Warren BA, Lee EA, Ding L:
 DR WPI: 2002-188744/24.
 DR P-PSDB: AAE18656.
 XX
 PT New human G-protein coupled receptor polypeptide for diagnosis,
 PT prevention and treatment of cell proliferative, neurological,
 PT cardiovascular, gastrointestinal, autoimmune/inflammatory, and
 PT metabolic disorders
 PS
 XX
 PS Claim 5, Page 149: 150pp: English.
 CC The invention relates to novel human G-protein coupled receptors
 CC (GCRRC) and their encoding polynucleotides. GCRRC is useful as an
 CC immunogen for preparing monoclonal and polyclonal antibodies. GCRRC is
 CC useful for diagnosing, treating and preventing a cell proliferative
 CC disorder (e.g., hepatitis, psoriasis, cancer), a neurological disorder
 CC (e.g., epilepsy, Alzheimer's disease, Pick's disease, Huntington's
 CC disease, Parkinson's disease), a cardiovascular disorder (e.g.,
 CC atherosclerosis, hypertension, myocardial infarction), gastrointestinal
 CC disorder (e.g., gastritis, cirrhosis, Crohn's disease), an autoimmune/
 CC inflammatory disorder (e.g., acquired immunodeficiency syndrome (AIDS),
 CC allergy, anaemia, asthma, rheumatoid arthritis), a metabolic disorder
 CC (e.g., diabetes, obesity, osteoporosis), and viral infections. GCRRC is
 CC useful in a number of drug screening techniques, and to analyse the
 CC proteome of a tissue or cell type. GCRRC is useful for creating knockin
 CC humanised animals or transgenic animals to model human diseases. In
 CC somatic or germline gene therapy, to generate a transcript image of a
 CC tissue or cell type, for detecting differences in the chromosomal
 CC location due to translocation, inversion, etc., among normal, carrier
 CC or affected individuals, and as hybridization probes for mapping
 CC naturally occurring genomic sequences. GCRRC is useful in Southern or
 CC northern analysis, dot blot or other membrane-based technologies, in PCR
 CC technologies, in distal, pin, multiformat enzyme linked immunosorbent
 CC (ELISA)-like assays, and in microarrays utilising fluids or tissues from
 CC patients to detect altered GCRRC expression. The present sequence is
 CC human GCRRC-17 cDNA.
 XX
 SQ .sequence 948 BP: 176 A; 279 C; 228 G; 265 T; 0 other;
 Query Match 85.5%; Score 810.4; DB 24; Length 948;
 Best Local Similarity 90.9%; Pred. No. 2.3e-217;
 Matches 862; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
 QY 1 ATGGAGAGATGGGTGACACAGTCTACACAGATGCTTCTCTGGGACATCTTTCC 60
 DB 1 ATGGAGAGATGGGTGACACAGTCTACACAGATGCTTCTCTGGGACATCTTTCC 60
 QY 61 CACAGCCAGACTGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120

| | |
|----|---|
| DE | Human G-protein coupled receptor 6 (GCR6-6) cDNA. |
| KW | Human; G-protein coupled receptor 6; cell proliferative disorder; |
| KW | arteriosclerosis; hepatitis; cancer; neurological disorder; epilepsy; |
| KW | Alzheimer's disease; Parkinson's disease; cardiovascular disorder; |
| KW | atherosclerosis; hypertension; myocardial infarction; peptic ulcer; |
| KW | gastrointestinal disorder; dysphagia; anorexia; autoimmune disorder; |
| KW | acquired immune deficiency syndrome; inflammatory disorder; infection; |
| KW | Addison's disease; allergy; Grave's disease; metabolic disorder; AIDS; |
| KW | diabetes; obesity; osteoporosis; gene therapy; GCR6-6; ss. |
| XX | |
| OS | Homo sapiens. |
| XH | |
| FH | Key Location/Qualifiers |
| FT | CDS 1..816 |
| FT | /tag= a |
| FT | /product= "Human GCR6-6 protein" |
| XX | |
| PN | WO200187937-A2. |
| XX | |
| PD | 22-NOV-2001. |
| XX | |
| PF | 17-MAY-2001; 2001WO-US16285. |
| XX | |
| PR | 18-MAY-2000; 2000US-205628P. |
| PR | 22-MAY-2000; 2000US-206222P. |
| PR | 25-MAY-2000; 2000US-207566P. |
| PR | 02-JUN-2000; 2000US-208834P. |
| PR | 02-JUN-2000; 2000US-208861P. |
| XX | |
| PA | (INCY-) INCYTE GENOMICS INC. |
| XX | |
| P1 | Patterson C, Lu DM, Thornton M, Lu Y, Tribouley CM, Graul R; |
| P1 | Khan FA, Gandhi AR, Walla NK, Nguyen DB, Yue H, Hatalla A; |
| P1 | Elliott VS, Lal P, Reddy R, Kallick DA, Tang TY, Au-Yang J; |
| XX | |
| DR | WPI: 2002-089844/12. |
| XX | |
| XR | P-PSDB; AAEL6175. |
| CC | |
| PT | Novel G-protein coupled receptors and polynucleotides useful for |
| PT | diagnosis, treatment and prevention of disorders of cell proliferation, |
| PT | neurological, cardiovascular, metabolic disorders and viral infections |
| XX | - |
| XX | |
| PS | Claim 5; Page 114; 115pp; English. |
| XX | |
| CC | The invention relates to human G-protein coupled receptor (GCR6) |
| CC | polypeptides and polynucleotides. GCR6 polypeptides are useful for |
| CC | screening compounds that modulate their activity. They are useful in |
| CC | the diagnosis, prevention and treatment of disorders which include |
| CC | cell proliferative disorders such as arteriosclerosis, hepatitis, |
| CC | myofibrosis, psoriasis and cancer including adenocarcinoma, leukaemia |
| CC | lymphoma, neurological disorders such as epilepsy, ischemic |
| CC | cerebrovascular disease, Alzheimer's disease, Pick's disease, dementia, |
| CC | Parkinson's disease, ataxias, multiple sclerosis, bacterial and viral |
| CC | meningitis, Creutzfeldt-Jakob disease, schizophrenic disorders, amnesia |
| CC | cardiovascular disorders such as arteriovenous fistula, atherosclerosis |
| CC | hypertension, vascular tumours, myocardial infarction, hypertensive |
| CC | heart disease, infective endocarditis, cardiomyopathy, myocarditis; |
| CC | gastrointestinal disorders such as dysphagia, peptic oesophagitis, |
| CC | emesis, anorexia, nausea, peptic ulcer, cholelithiasis, diarrhoea, |
| CC | constipation, acquired immune deficiency syndrome (AIDS), hepatic |
| CC | encephalopathy; autoimmune/inflammatory disorders such as Addison's |
| CC | disease, allergies, spondylitis, amyloidosis, anaemia, asthma, contact |
| CC | dermatitis, Crohn's disease, diabetes mellitus, Goodpasture's syndrome, |
| CC | emphysema, Grave's disease, gout, multiple sclerosis, rheumatoid |
| CC | arthritis, systemic lupus erythematosus, uveitis, viral, bacterial, |
| CC | fungal, parasitic, protozoal and helminthic infections and truma; |
| CC | metabolic disorders such as diabetes, obesity and osteoporosis; and |
| CC | viral infections such as infection caused by viral agent classified as |
| CC | adenovirus, arenavirus, bunyavirus. Polynucleotides of the invention |
| CC | are used as probes for assessing toxicity of test compounds. They are |
| CC | also used in gene therapy. The present sequence is human G-protein |

| CC | coupled receptor 6 (GOREC-6) cDNA. |
|----|---|
| XX | Sequence 965 BP; 181 A; 270 C; 244 G; 270 T; 0 other: |
| SO | Query Match 70.9%; Score 672.4; DB 24; Length 965; Best Local Similarity 86.0%; Pred. No. 1,3e-178; Matches 815; Conservative 0; Mismatches 1; Indels 132; Gaps |
| OY | 1 ATGGAGATATGGGAGAACCACTCCCTACACAGATGGCTTCTCCCTGGGCACTTTTCC 60 |
| DB | 1 ATGGAGATATGGGAGAACCACTCCCTACACAGATGGCTTCTCCCTGGGCACTTTTTC 60 |
| OY | 61 CACAGCCAGACTGACCTTGTCTCTCTCTGCACTTAATGATGCTTTCACAGTGGCCCTC 120 |
| DB | 61 CACAGCCAGACTGACCTTGTCTCTCTCTGCACTTAATGATGCTTTCACAGTGGCCCTC 120 |
| OY | 121 TGTGGGAATGTCTCCCTCATCTTCTCTCATCTGACAGCTGGACTTCACACCCCATG 180 |
| DB | 121 TGTGGGAATGTCTCCCTCATCTTCTCTCATCTGACAGCTGGACTTCACACCCCATG 180 |
| OY | 181 TACCTTCTCCAGCAGCTCTCCCTCATGAGACCTCATGTTGGTCTTAAACATTTGSCA 240 |
| DB | 181 TACTTCTTCCACCCAGCTCTCCCTCATGAGACCTCATGTTGGTCTTAAACATTTGCCA 240 |
| OY | 241 AAGATGGCAGCCAACTTCTGTGCGAGAGAAATCATCTCTTGTGGGCTGTGGCATA 300 |
| DB | 241 AAGATGGCAGCCAACTTCTGTGCGAGAGAAATCATCTCTTGTGGGCTGTGGCATA 300 |
| OY | 301 CAATTTGGCTTTTGTCTCTCTGTGGGATCTGAGGGGCTCTTGGCACTCATGGCT 360 |
| DB | 301 CAATTTGGCTTTTGTCTCTCTGTGGGATCTGAGGGGCTCTTGGCACTCATGGCT 360 |
| OY | 361 TATGACCACTACGTGGGCGTTATGACACCACTTCACATCTCCATCCATGAATCAGAG 420 |
| DB | 361 TATGACCACTACGTGGGCGTTATGACACCACTTCACATCTCCATGAATCAGAG 420 |
| OY | 421 GTCTCTCCAGATTACTGGGAGCTCCTGGGCTTTGGGATAATAGATGAGATTCAG 480 |
| DB | 421 GTCTCTCCAGATTACTGGGAGCTCCTGGGCTTTGGGATAATAGATGAGATTCAG 480 |
| OY | 481 ATGTGGCAGCATGGGCTTACCTTACTGTGGGCTCAAGGAGCGGTGATCATCTTTTCTG 540 |
| DB | 481 ATGTGGCAGCATGGGCTTACCTTACTGTGGGCTCAAGGAGCGGTGATCATCTTTTCT 540 |
| OY | 541 GAGTACAACTTTATTGAGCTGGCTGTGACAGACACTTCCCTTTTGACACCCCTCTC 600 |
| DB | 541 GAGTACAACTTTATTGAGCTGGCTGTGACAGACACTTCCCTTTTGACACCCCTCTC 600 |
| OY | 601 TTTGCTGTGTGTCTTCATGCTTCTCTTCCCTTCTTCATCATCATGAGCCTCTATGCT 660 |
| DB | 601 TTTGCTGTGTGTCTTCATGCTTCTCTTCCCTTCTTCATCATCATGAGCCTCTATGCT 660 |
| OY | 661 TGCATCTAGGGGCTGTGCTCCGAATACGCTCTGCTCAGGCGTGGAAAAAGCCCTGGCC 720 |
| DB | 661 TGCATCTAGGGGCTGTGCTCCGAATACGCTCTGCTCAGGCGTGGAAAAAGCCCTGGCC 720 |
| OY | 721 ACCGTCTCTCCACCTTAACAGCTGTCAACCTCTCTTATATGGGGAGCCATGTTTCATG 780 |
| DB | 721 ACCGTCTCTCCACCTTAACAGCTGTCAACCTCTCTTATATGGGGAGCCATGTTTCATG 780 |
| OY | 781 CTGAGGCGCTAGGCGCTACCGGGCCCTTAGCCATCACAAGGTGGGCTTATCTTCTACACA 840 |
| DB | 781 CTGAGGCGCTAGGCGCTACCGGGCCCTTAGCCATCACAAGGTGGGCTTATCTTCTACACA 840 |
| OY | 841 GTCTTACTCCCATGTGAACCCCTCATTTACAGCTTGTGAGAAATGGGAGGTGATGGGG 900 |
| DB | 841 GTCTTACTCCCATGTGAACCCCTCATTTACAGCTTGTGAGAAATGGGAGGTGATGGGG 900 |
| OY | 901 GCATGAGAAAGGGGCTGAGACCGGTGAGGATTTGGCAGCCAGCACTGA 948 |
| DB | 901 GCATGAGAAAGGGGCTGAGACCGGTGAGGATTTGGCAGCCAGCACTGA 948 |
| OY | 948 GCACTTGTGAGAAAGGGGCTGAGACCGGTGAGGATTTGGCAGCCAGCACTGA 986 |
| DB | 948 GCACTTGTGAGAAAGGGGCTGAGACCGGTGAGGATTTGGCAGCCAGCACTGA 986 |

```

RESULT 13
AA51179
ID AAA51179 standard; cDNA; 669 BP.
XX
AC AAA51179;
XX
DT 26-SEP-2000 (first entry)
XX
DE Murine olfactory receptor ligand-binding region cDNA 5.
XX
KW Degenerate primer; ligand-binding region; olfactory receptor;
KW Transmembrane domain; N-terminal; plasma membrane; translocation domain;
KW human rhodopsin receptor; odorant; toxicity; olfactory response; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 1..669
FT /tag=a
FT /product= Olfactory_receptor_ligand_binding_protein
FT /partial
XX
PN MO200035274-A1.
XX
PD 22-JUN-2000.
XX
PF 17-DEC-1999; 99WO-US30221.
XX
PR 17-DEC-1998; 98US-0112605.
XX
PA (UYUO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Reed RR, Krautwurst D, Yau KW;
XX
DR WPI: 2000-431471/37.
XX
DR P-PSDB: AAY96666.
XX
PT Primer pair for amplifying olfactory receptor nucleic acid, useful e.g.
PT for producing receptor libraries used in e.g. screening odorants for
XX toxicity
XX
PS Claim 18; Page 31-32; 61pp; English.
XX
CC AAA51175-93 encode murine odorant/ligand binding regions of olfactory
CC receptors. They were amplified using degenerate primers shown in
CC AAA51163-64.
CC The primers can be used to generate a library of OR-LBR comprising the
CC transmembrane domains (TMD) II-VI, III-VI, II-VII or III-VII. Chimeric
CC nucleic acids encoding a 7-TMD protein comprise a sequence encoding an
CC N-terminal plasma membrane translocation domain (especially the sequence
CC shown in AAY96658), a first TMD and the OR-LBR. The translocation domain
CC was initially derived from the N-terminus of the human rhodopsin
CC receptor. LBR produced by amplification with the primers are useful for
CC generating new odorants, to screen for toxicity or therapeutic activity
CC in odorants, and altering an animal's olfactory response (claimed).
XX
SQ Sequence 669-BP; 124 A; 190 C; 148 G; 207 T; 0 other;
XX
Query Match 55.5%; Score 526.6; DB 21; Length 669;
Best Local Similarity 86.7%; Pred. No. 9.6e-138;
Matches 580; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
XX
QY 193 AGCGAGCTCTCCCTGACCTGATGTTGGTGTATACATTTGCCAAAGATGCGACGC 252
DB 1 AGTCAGCTCTCCCTGACCTGATGTTGGTGTATACATTTGCCAAAGATGCGACGC 60
QY 253 AACTTCCTGTCTGGGAGAGATGCTCTTTGGGCTGTGGCATACAAATGGCTTT 312
DB 61 AACTTCCTGTCTGGGAGAGATGCTCTTTGGGCTGTGGCATACAAATGGCTTT 120
QY 313 TTGTGCTCTCTTGGGAGATGAGGGGCTCTTGGGAGCATGAGGCTTATGACCATAC 372
DB 121 TTGTGCTCTCTTGGGAGATGAGGGGCTCTTGGGAGCATGAGGCTTATGATGCTAT 180

```

```

QY 373 GTGGCCGTTAGCCACCACTTCACTATCCATCTTCATGATAGAGGGCTGTCTCCAG 432
DB 181 GTGGCCATTTAGCCACCACTTCACTATCCATCTTCATGATAGAGGGCTGTCTCCAG 240
QY 433 ATTACTGGGAGGCTCCGTGGGCTTTGGGATATAGATGAGTATTCAGATGCTGGCAGCC 492
DB 241 ATTACTGGGAGGCTCCGTGGGCTTTGGGATATAGATGAGTATTCAGATGCTGGCAGCC 300
QY 493 ATGGGCTTACCTTACTGTGTGCTCAAGAGGCTGATCATTCTTGTGTGAGGTACAAAGCT 552
DB 301 ATGAGGCTTCCCTTACTGTGTGCTCAAGAGGCTGATATGATATCATTCTTGTGTGAGGTGCGGCT 360
QY 553 TTATTGAAGCTGGCCGTGGGCTGAGCACTTCCCTTTTGAAGCCCTGCTTGTGTGTGT 612
DB 361 TTACTGAAGCTGGCCGTGGGCTGAGCACTTCCCTTTTGAAGCCCTGCTTGTGTGTGT 420
QY 613 GTCTTCATGCTTCTCCCTTCCCTTCTTCATCATCATGAGCCCTCATGCTTGCATCCTAGGG 672
DB 421 GTCTTCATGCTTCTCCCTTCTTCATCATCATGAGCCCTCATGCTTGCATCCTAGGG 480
QY 673 GCTGTGCTCCGAATACGCTGTCTGAGGCTGGAAAAAGCCCTGGCCACTGTCTCTCC 732
DB 481 GCTGTGCTCCGAATACGCTGTCTGAGGCTGGAAAAAGCCCTGGCCACTGTCTCTCC 540
QY 733 CACCTAAGAGCTGTGACCCCTTCTATGAGGAGGAGCATGTTATCTTACTGAGGCTTAGG 792
DB 541 CACCTAAGAGCTGTGACCCCTTCTATGAGGAGGAGCATGTTATCTTACTGAGGCTTAGG 600
QY 793 CGCTACCGGGGCCCTAGGCATGACAAAGTGGCTCTATCTTATACAGAGTCTTACTCC 852
DB 601 CGATATTCGGGCTCTTACCATGACAAAGTGTCTCATCTTCTTACAGAGTCTTACTCTCT 660
QY 853 ATGCTGAC 861
DB 661 ATGCTGAC 669
XX
RESULT 14
AA51183
ID AAA51183 standard; cDNA; 669 BP.
XX
AC AAA51183;
XX
DT 26-SEP-2000 (first entry)
XX
DE Murine olfactory receptor ligand-binding region cDNA 9.
XX
KW Degenerate primer; ligand-binding region; olfactory receptor;
KW Transmembrane domain; N-terminal; plasma membrane; translocation domain;
KW human rhodopsin receptor; odorant; toxicity; olfactory response; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 1..669
FT /tag=a
FT /product= Olfactory_receptor_ligand_binding_protein
FT /partial
XX
PN MO200035274-A1.
XX
PD 22-JUN-2000.
XX
PF 17-DEC-1999; 99WO-US30221.
XX
PR 17-DEC-1998; 98US-0112605.
XX
PA (UYUO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Reed RR, Krautwurst D, Yau KW;
XX
DR WPI: 2000-431471/37.

```


QY 313 TTGTCCTCTGTGGAGATCGAGGGCTTGTGCTGGAGCTCATGGCTTATGACCCTAC 372
 Db 121 TTGTCCTCTGTGGAGATCGAGGGCTTGTGCTGGAGCTCATGGCTTATGATGCTAT 180
 QY 373 GTGGCCGTAGGCAACCCACTCATGCCATCCCATGAAATCAGAGGCTGTCTCCAG 432
 Db 181 GTGGCCATTAACCCACCTTCACTCCATCCATTCATGAGGCCAAAGGTGTCTCCAG 240
 QY 433 ATTAAGGAGCTCTCGGGCTTTGGATTAATAGATGAGTGAATCAGATGAGGACACC 492
 Db 241 ATTGCTGAATCTCTCGGGCTTTGGAGATCCTGATGGAATTAATTCAGATGAGGACCC 300
 QY 493 ATGGGCTTACCTTACTGTGCTCAAGAGCGTGATCACTTTTCTGTGAGGTACAAGCT 552
 Db 301 ATGAGCCCTGCCCTACTGTGGCTCAAGGATTAATGATCACTTCTGTGAGGTGCCGCT 360
 QY 553 TTATGAGCTGGGCTGTGAGACACTTCCCTTTTGACACCCCTCTTGTGCTGTGT 612
 Db 361 TTACTGAAGCTGGCTGTGAGACACCTCCCTTTCGACACCCCTGCTTGTGCTGTGT 420
 QY 613 GTCTTCATGCTTCTCCCTTCCCTTCCATCATGAGGCTCTCATGCTTGCATCCTAGG 672
 Db 421 GTCTTATGCTGCTTCTCTCTTCTCGATCATGTGACTTCTTATGCTGCACTTGGGG 480
 QY 673 GCTGTGCTCGGAATACGCTGTGCTGAGGCTGAAAAAGCCCTGGCCACCTGCTCTCC 732
 Db 481 ACTGTGCTCGGATGACACTGTGCCAGTCCCGAAAAAGGCCCTGGCCACTGTCTCC 540
 QY 733 CACCTAAGAGCTGTACCCCTTCTATGAGGAGCATGTTCATGTACCTGAGGCTTAGG 792
 Db 541 CACCTGACAGCTGTCTCTTCTACGAGGAGCAATGTTCATCTACCTGAGGCGCAAG 600
 QY 793 CGCTACCGGGCCCTAGCCATGACAAGSTGSCCTATCTCTACACAGTCTTACTCCC 852
 Db 601 CGATATGCGGCTCTCTAGCCATGACAAGTGTCTCAATCTCTACACAGTCTTACTCCT 660
 QY 853 ATGCTGAAC 861
 Db 661 ATGCTCAAC 669

Search completed: May 9, 2003, 03:37:26
 Job time : 272 secs

THIS PAGE BLANK (USPTO)

Class: BAC ends
High quality sequence stop: 398.
Location/Qualifiers
1. 398

FEATURES
SOURCE
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-2250 Col=7 Row=0"
/sex="male"
/note="Organ: sperm; Vector: pBeloBAC11; BAC clones in E-coli DH10B"

BASE COUNT 121 a 103 c 98 g 75 t 1 others
ORIGIN

Query Match 40.1%; Score 380.6; DB 17; Length 398;
Best Local Similarity 97.5%; Pred. No. 1.5e-87;
Matches 386; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

151 TACCTGACGCTGGACCTGACACCCCATGATCTCTCTGACCCAGCTCTCCCTCATG 210
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB TACGTGACGCTGGACCTGACACCCCATGATCTCTCTGACCCAGCTCTCCCTCATG 337
211 GACCTCATGCTGCTGTGATACATGTGGCAAGATGGACCACTTCTCTGCGCAG 270
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB GACATCATGCTGCTGTGATACATGTGGCAAGATGGACCACTTCTCTGCGCAG 277
271 AAGTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 330
276 AAGTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 217
331 TCTGAGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 390
216 TCTGAGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 157
391 CTTCACATCTCCATCCATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 450
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB CTTCACATCTCCATCCATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 97
451 GCCTTGGGATATAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 510
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB GCCTTGGGATATAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 37
511 GGTCTAGAGAGCGGTGATCACTTTTCTGTGAGTA 546
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB GGTCTAGAGAGCGGTGATCACTTTTCTGTGAGTA 1

RESULT 2
BB538338
LOCUS BB538338 664 bp mRNA linear EST 26-OCT-2001
DEFINITION BB538338 RIKEN full-length enriched, 0 day neonate eyeball Mus
ACCESSION BB538338
VERSION BB538338.2 GI:16446878
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
1 (bases 1 to 664)
Atakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
'M., Koyama,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,K., Sakai,K., Sano,H., Sasaki,
'D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Atakawa,T., et al. 2001)
Unpublished (2001)
On Jul 31, 2000 this sequence version replaced gi:9593838.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayashi,N., Sugahara,Y., Shibata,K., Itoh
'M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
'S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multichannel sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
'Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
'Kondo,S., Shingawa,A., Saito,T., Kiyosawa,H., Yamana,I., Aizawa
'K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

FEATURES

SOURCE

Location/Qualifiers
1. 664

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E130011E23"
/sex="male"
/note="Organ: sperm; Vector: pBeloBAC11; BAC clones in E-coli DH10B"

BASE COUNT 141 a 176 c 142 g 204 t 1 others
ORIGIN

Query Match 34.6%; Score 327.8; DB 10; Length 664;
Best Local Similarity 86.2%; Pred. No. 8.6e-74;
Matches 362; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

529 CACTTTCTCTGCTGATACAGCTTTATTAAGCTGCTGTGACGACACTCTCTTTT 588
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB CACTTTCTCTGATACAGCTTTATTAAGCTGCTGTGACGACACTCTCTTTT 60
589 GACACCTCTCTCTGCTGCTGCTTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 648
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

| | | | |
|-------------|---|--|------|
| Db | 61 | GACACCCTGCCTTGGCTTGCTGTCGTCTNTAATAGTCGTTCTTCCTTCGCATCAATTGG | 120 |
| OY | 649 | GCCTCCATTCGCTTGATCCTCTAGGGCGCTGCTCCGAATAACCTCTGCTCAGCCCTGGAAA | 708 |
| Db | 121 | ACTCTTAATGGCTGGCATCTTGTGGGGGCTGTGCCTCCATATGACCTCTGCCACGCCGAA | 180 |
| OY | 709 | AAAGCCCTGGCCACCATCGTCTCCGCCACCTAACAGCTGTACACCTCTCTTATATGGGACAG | 768 |
| Db | 181 | AAGGCCCTGGCCACCTGTGTTCCTCCACCTGCAGCTGTCTCTCTTCTTAGGGGACAGA | 240 |
| OY | 769 | ATGTTCAATGTAACCTGAGGCTTAGGCGCTACCGGGGCCCTTAGCCATAGACAAGGTGGCTCT | 828 |
| Db | 241 | ATGTTCAATGTAACCTGAGGCGCAAGCGGATATGCGCTCTAGCACATGACAATAATTTGTCCA | 300 |
| OY | 829 | ATCTCTCAACAGTCCCTTACCTCCATCCATGGTGMAACCCCTCATTTACAGCTTGAAGAAATGG | 888 |
| Db | 301 | ATCTCTCAACAGTCTTACTCTTATGCTCAACCCCTCATTTAATAGCTTAGGAACAGG | 360 |
| OY | 889 | GAGGTGATGGGGGCGACTGAGGAAGGGGGCTGACCGCTGCAGGATTGGCAGCCAGCACTGA | 948 |
| Db | 361 | GAGGTGATGGGGGCGACTGAGGAAGGAGACTGGACCGCTGCAGGATTGGCAGCCAGCAATTGA | 420 |
| RESULT 3 | | | |
| AKO16560 | | | |
| LOCUS | | | |
| DEFINITION | AKO16560 | 3063 bp | mRNA |
| DESCRIPTION | Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493244IH21; similar to T1 OLFACTORY RECEPTOR, full insert sequence. | | |
| VERSION | AKO16560 | | |
| KEYWORDS | AKO16560.1 GI:12855357 | | |
| SOURCE | HTC: CAP trapper. Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone_1ib:RIKEN full-length enriched mouse CDNA library clone:493244IH21. | | |
| ORGANISM | Mus musculus | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| REFERENCE | | | |
| AUTHORS | 1 Carninci,P. and Hayashizaki,Y. | | |
| TITLE | High-efficiency full-length cDNA cloning | | |
| JOURNAL | Meth. Enzymol. 303, 19-44 (1999) | | |
| PUBMED | 99279253 | | |
| PUBMED | 10349636 | | |
| REFERENCE | | | |
| AUTHORS | 2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. | | |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes | | |
| JOURNAL | Genome Res. 10 (10), 1617-1630 (2000) | | |
| PUBMED | 20499374 | | |
| PUBMED | 11042159 | | |
| REFERENCE | | | |
| AUTHORS | 3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Komono,H., Akiyama,T., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishie,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Katohki,M., Yoneeda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y. | | |
| TITLE | RIKEN integrated sequence analysis (RISn) system--384-format sequencing pipeline with 384 multicapillary sequencer | | |
| JOURNAL | Genome Res. 10 (11), 1757-1771 (2000) | | |
| PUBMED | 20530913 | | |
| PUBMED | 11076861 | | |
| REFERENCE | | | |
| AUTHORS | 4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Komono,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gotohori,T., Bono,H., Kusukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavault,T., KleiSchmann,W., Gaasterland,Y., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuno,Y., Nikaido,I., Pesole,G., | | |

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Quackenbush, J., Schriml, L.M., Straubli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Caninelli, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustinchik, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamlay, M., Lee, N.H., Lyons, P.,
 Mochizuki, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Saeki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Scorch, K.F., Suzuki, H.,
 Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
 Wyszawski-Borts, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlsuki, S.,
 and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

5 (bases 1 to 3063)
 Adachi, J., Alzawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arikawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
 Caninelli, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
 Hume, D., Imotoani, K., Ishii, Y., Itoh, M., Izawa, M., Katsukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koyas, S.,
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
 Nunazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D.,
 Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T.,
 Soejima, Y., Suzuki, H., Tagami, Y., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamakake, I.,
 Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and
 Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration and Chemical Research, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer
 [5'-GAGAGAGACAGATCCAGTAAATTAATTAATGATCCGCCCCCCCCC 3']. cDNA was
 prepared by using trehalose thermo-activated reverse transcriptase
 and subsequently enriched for full-length by cap-trapper. cDNA went
 through one round of normalization to Rot = 10.0 and subtraction to
 Rot = 100.0. Second strand cDNA was prepared with the primer
 adapter of sequence 15'
 GAGAGAGACATTCCTCAGTAAATTAATTAATGATCCGCCCCCCCCC 3'. cDNA was cleaved
 with BamHI and XhoI. cDNA of size compressed between 3 and 7 kb was
 selected before cloning. Vector: a modified plasmidscript KSI(+) after
 bulk excision from lambdaB FLC I. Cloning sites, 5' end: SalI, 3'
 end: BamHI. Host: DH10B.

FEATURES
SOURCE

Location/Qualifiers
 1..3063
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="FANTOM.DB:4932441H21"
 /db_xref="MGD:MGI:180765"
 /db_xref="taxon:10090"
 /clone="4932441H21"
 /sex="male"
 /tissue_type="testis"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 438..1388
 /note="data source:SPTR, source key:09EPF8, evidence:ISS
 putative
 similar to T1 OLFACTORY RECEPTOR"

CDS

```

/codon_start=1
/protein_id="BAB30304.1"
/db_xref="GI:12853358"
/translation="MEPNSSTLESFGLIDSGSPPELLCATVTLMLISNGL
LILVITVDRLHVPMTLLRLSLIDLTSTVPTVVDLRLDRTISFEQALOLF
SAATLIGAEELLAEMADRYVAICHPLNMIFMSKARLMAISWILASALHT
VYTMHPFCMSOEIRHLICEVPLKLAADSQVLEMYVTVGIFLLPLSLITTSY
SLITFVLHMPDSNEGRKALMTCSSHLTVGMFYGGATFMYVLPSSSEHSPKODNIISV
FVITVPALNPFLYSLRNKEVIGAVRVRIGRHLLPAHATV"
polya_signal
3038..3043
/motif="putative"
polya_site
3063
/note="putative"
BASE COUNT      879 a      612 c      696 g      876 t
ORIGIN

```

```

Query Match      31.3%; Score 297; DB 11; Length 3063;
Best Local Similarity 58.2%; Pred. No. 1.7e-65;
Matches 522; Conservative 0; Mismatches 375; Indels 0; Gaps 0;

```

```

23 CCTACACAGATGGCTCTTCCTTGGGCACTTTTCCACAGCCAGACTGACCTTGTCC 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 457 CCTGGAAAGTGGATTTATCTTGGTGGGATTCGATGGCAGTGGCTCTCTGAACTGC 516

QY 83 TCTTCTCGAGTTAGTGGTCTTCAAGTGGCCCTGTGGAAATGCTCCTCACT 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 517 TCTGTGCACAGTTAACACCTGTACATGTTGGCAGTACAGCAATGACTTCTACTCC 576

QY 143 TCCCATCTACCTGGAGCGTGGACTCACACCCCATGTAATCTTCTCTCCAGCCGCT 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577 TGTGATCAGACATGATGAGCCCGGCTTCACTGACCAATGTAATCTCTCTGAGCACTGT 636

QY 203 CCCATGATGACCTGATGTTGGTCTGTAACTTGTCCAAAGATGGACCAACTTCTGT 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 637 CTCATCTAGACTCTCTCTCACTGATGTAATCCCAACACGTGTGGATTTTCTGC 696

QY 263 CTGCAGAGATCCATCTCTTGTGGGCTGTGCATACAAATGCGTTTGTCTCTC 322
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 697 TCAGAGACACACCATATCTTGTAGGAGATGCGCTTCAATTTCTTCACCAATGACAT 756

QY 323 TTGTGGATCTGAGGGGCTGTGGTGGGACATGAGCTTATGACACATGAGCGGCTTA 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 757 TGGGTGTGACAGAGAGCTCTCTGCGCTTCAATGAGCTTATGATAGTATGCGCATTT 816

QY 383 GCCACCCACTTCACTATCCATCTCATGAATCAGAGGGTCTGTCTCCAGTTACTGGA 442
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 817 GTCAATCTCTTAACATCAATGATCTTCAATGATCCAAAGGCTGCAAGCTCATGTGCGCA 876

QY 443 GCTCTGGGCTTTGGGATATATGATGAGATTCAGATGATGCGCAGCCATGGGCTTAC 502
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 877 TATCATGATCTTACGATCTTCTAGTGCCCTAGGTCACAGTGTACCAATGACATCTCC 936

QY 503 CTATACCTGGGCTCAAGAGCGTGTGACTTTTCTGTAGAGTACAACTTATTTGAAGC 562
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 937 CTTTTCATGATGCCAGAAATCAGACACGCTGTCTGTAGAGTTCCTCATTTGTTGAAT 996

QY 563 TGGCCTGTGACAGACCTTCCCTTTTGTACACCCCTCTCTTGTGCTGTCTTCAATGC 622
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 997 TGGCTTGTGACAGACATCTCAATATGAGCTCATGAGTTATGTGACAGAGATGATTC 1056

QY 623 TTTCCTTCCCTTCTTCATCATCATGAGGCTCCTATGCTTGAATCTAGGGGTGTGCTCC 682
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1057 TATTCCTCTCCCTCTCTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1116

QY 683 GAATACGCTGTGCTGAGCGCTGGAAGAAAGCGCTGACCTGCTCCCACTCACTAAGAG 742
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1117 ACATCCCTTCAATATAGGCGAGAAAGAAAGCGCTGTGACCTGCTTCCCACTGATCTG 1176

QY 743 CTGTACACCTCTCTTCTTATAGGGGAGCCATGTTATCTACCTAGAGGCTTACCGCGG 802
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1177 TGGTTGGAGATGTTCTATAGGGGGGTGCCACTTTCATGTATGTGTGCCAGATTCTTCCACA 1236

```

```

QY 803 CCCCTAGCCATGACAGGTGGCTCTATCTTCTACACAGTCCCTTACCCATCGTGAACC 862
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1237 GTCTTAGACAGACAAATCATCTGTGTCTATACAAATTTGCACACCGCTCTGAACC 1296

QY 863 CCTCATTTACACCTTGAGATGGAGGTGATGGGGCACTGAGGAAGGCGCTGG 919
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1297 CCTCATTTACACCTTGAGAAATAGAGGTGATTTGAGACTGTAGAAAGTACTGG 1353

RESULT 4
AK017036
LOCUS
DEFINITION
1394 bp mRNA linear HTC 19-JAN-2002
Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:493343E02:similar to T1 OLFATORY RECEPTOR, full
insert sequence.
AK017036
AK017036
VERSION
AK017036.1 GI:12856091
KEYWORDS
HTC; CAP-trapper.
MUS musculus (Strain:C57BL/6J) adult male testis cDNA to mRNA,
clone:11b:RIKEN full-length enriched mouse cDNA library
SOURCE
MUS musculus
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Eukaryota.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
REFERENCE
AUTHORS
JOURNAL
MEDLINE
PUBMED
TITLE
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
MEDLINE
PUBMED
TITLE
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arai, K., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, K.,
Saito, T., Okazaki, Y., Gotojori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuenl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J.,
Schrimal, L. M., Staudt, J., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bona, M. F., Brownstein, M. J., Bult, C., Fletcher, C.,
Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombauer, P., Nordone, P.,
Ringo, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyono, O., K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
JOURNAL

```



```

/clone="UNGCM0562E19"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/sex="Male"
/note="Vector: PMD42nv: Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1147321149b1ar12972.1), a copy-number
inducible derivative of plasmid RL. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

```

BASE COUNT      124 a      195 c      129 g      184 t
ORIGIN

```

```

Query Match      26.3%; Score 249.4; DB 17; Length 632;
Best Local Similarity 63.9%; Pred. No. 1.5e-53;
Matches 393; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

```

```

OY 166 CTTCACACCCCATGCTACTTCTTCTCAGCAGCTCTCCCTCATGACCTCATGTC 225
    || || || || || || || || || || || || || || || || || || || ||
DB 19 CTCACATACCAATGCTACTTTTATACAGCCAGCTGCCCATGACATGATGATATT 78
    || || || || || || || || || || || || || || || || || || || ||
OY 226 TGTAACTATGTCGCAAGAGTGACCACTTCTGTCGTGAGAGAGTCCATCTT 285
    || || || || || || || || || || || || || || || || || || || ||
DB 79 TCTGTCACTGTGCGCAAGATGCTATGACAGCTCTCGGGAGCCACAAAGATCTAGCT 138
    || || || || || || || || || || || || || || || || || || || ||
OY 286 GTGGGCTGTGGCATACAAATGGCTTTTGTCTCTCTGTGGAGTCTGAGGGCTCTTG 345
    || || || || || || || || || || || || || || || || || || || ||
DB 139 GCTGCTGTGGAGTGAAGATGTTCTCTACTGACATTAGAGAGGTCAGAAATATTCTT 198
    || || || || || || || || || || || || || || || || || || || ||
OY 346 CTGGGACTCATGCTTATGACCACTACGTGGCCCTTAGCCACCCACTTCACTATCCATC 405
    || || || || || || || || || || || || || || || || || || || ||
DB 199 CTTCCTGCAATGCTCTTATGACCACTATGTCGCACTGTATCCATCCAGATATCTGTC 258
    || || || || || || || || || || || || || || || || || || || ||
OY 406 CTCAATGATCAGAGGCTGCTGCCAGATTAATCTGGAGCTCTCGGCTTTGGATATA 465
    || || || || || || || || || || || || || || || || || || || ||
DB 259 CTCAATGATCAGAGGCTGCTGCCAGATTAATCTGGAGCTCTCGGCTTTGGATATA 318
    || || || || || || || || || || || || || || || || || || || ||
OY 466 GATGAGTGAATTCAGATGATGAGGAGGCTTACTTACTGTGCTCAAGAGAGCTG 525
    || || || || || || || || || || || || || || || || || || || ||
DB 319 GATGAGTGAATTCAGATGATGAGGAGGCTTACTTACTGTGCTCAAGAGAGCTG 378
    || || || || || || || || || || || || || || || || || || || ||
OY 526 GATGAGTGAATTCAGATGATGAGGAGGCTTACTTACTGTGCTCAAGAGAGCTG 585
    || || || || || || || || || || || || || || || || || || || ||
DB 379 CATGACTTCTGCTGAGGCTGCTGCTGAGCAAGAGCTCTGCTGAGCACTGCTGCTC 438
    || || || || || || || || || || || || || || || || || || || ||
OY 586 TTTGACACCCCTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
    || || || || || || || || || || || || || || || || || || || ||
DB 439 TATGAGACCCCTCATGTTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCT 498
    || || || || || || || || || || || || || || || || || || || ||
OY 646 ATGGCTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
    || || || || || || || || || || || || || || || || || || || ||
DB 499 TCAGGCTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 558
    || || || || || || || || || || || || || || || || || || || ||
OY 706 AAAAAGCCCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
    || || || || || || || || || || || || || || || || || || || ||
DB 559 AAGAAGGCTCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 617
    || || || || || || || || || || || || || || || || || || || ||
OY 766 GCCATGTTCAATGATC 780
    || || || || || || || || || || || || || || || || || || || ||
DB 618 GCTATCTATACCTAC 632

```

```

RESULT 7
AC067024/c 650 bp DNA linear GSS 03-NOV-2001
LOCUS Pan troglodytes DNA, clone: PTB-056004.R, genomic survey sequence.
AC067024
DEFINITION AC067024.1 GI:16618826
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Toto, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of library PTB
Unpublished
2 (bases 1 to 650)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Toto, Y., Watanabe, H. and Sakaki, Y.
DIRECT SUBMISSION
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpansegsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
LIBRARY
Sequencing: M13Rev
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. 650
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-056004.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

```

```

BASE COUNT      206 a      142 c      178 g      124 t
ORIGIN

```

```

Query Match      23.4%; Score 221.4; DB 17; Length 650;
Best Local Similarity 62.6%; Pred. No. 2.5e-46;
Matches 345; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

```

```

OY 143 TCTCATCTACCTGAGCGCTGACCTTCACACCCCATGCTTCTTCTGACGACTCT 202
    || || || || || || || || || || || || || || || || || || || ||
DB 608 TCTTATTCACAAAGAGATTTGCCCTTCATACACCATGCTCTCCCTGACGACTCTT 549
    || || || || || || || || || || || || || || || || || || || ||
OY 203 CCTCATGAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 262
    || || || || || || || || || || || || || || || || || || || ||
DB 548 CCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 489
    || || || || || || || || || || || || || || || || || || || ||
OY 263 CTGACAGAGATCATCTCTTGTGAGGCTGAGCATACAAATGGCTTTTGTCTCTC 322
    || || || || || || || || || || || || || || || || || || || ||
DB 488 TGATGATGAGGAGGCTCTTGTGAGGCTGAGCATACAAATGGCTTTTGTCTCTC 429
    || || || || || || || || || || || || || || || || || || || ||
OY 323 TTTGAGATGAGGAGGCTCTTGTGAGGCTGAGCATACAAATGGCTTTTGTCTCTC 382
    || || || || || || || || || || || || || || || || || || || ||
DB 428 TTTGAGATGAGGAGGCTCTTGTGAGGCTGAGCATACAAATGGCTTTTGTCTCTC 369
    || || || || || || || || || || || || || || || || || || || ||
OY 383 GCCACCATCTACCTATCCCATCTCATGATGATGATGATGATGATGATGATGATG 442
    || || || || || || || || || || || || || || || || || || || ||
DB 368 GCAACCTCTGATATACCTGCTGATGATGATGATGATGATGATGATGATGATG 309
    || || || || || || || || || || || || || || || || || || || ||
OY 443 GCTCTGGGCTTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 502

```

```

Db 308 GTTCCGTGGTGGGGCTCTTTGGATGGCTTCCCTCAACCCCATCCATCAGCTTTTC 249
QY 503 CTTCATGTGGCTCAGAGAGCGTGTGATCATTTTCTGTAGTACAGCTTTATAGC 562
Db 248 CCTTCCTCAATTTCCCGAGATTAAACCATCTTCTGTGAGGCGACGACTCTGAGT 189
QY 563 TGGCCTGTGACAGACATCTCCCTTTTGTACACCCCTCTTGTGTGCTGTTCATGC 622
Db 188 TGGATGTGACAGACAGCCCTCTATGAGACAGTATGATGTGCTGTGTTTGTATGC 129
QY 623 TTCTCTTCCCTTCCATCATCATCATGCGCTCTCTATGCTTGCATCTTACGGGCTGTGCTCC 682
Db 128 TGTGATTTCTTCTTCTGTATGCTCTTCTCTATGCGCCGATCTCTGATACATGCTAGT 69
QY 683 GAATACGCTCT 693
Db 68 GCATGAGCTCT 58

```

```

RESULT 8
A0558499/c 592 bp DNA linear GSS 29-MAY-1999
LOCUS HS_2082_B1.A05_MR CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate=2082 Col=9 Row=B, DNA sequence.
ACCESSION A0558499
VERSION A0558499.1 GI:4918231
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
JOURNAL High Throughput Sequencing Center
MEDLINE Contact: Mahairas GG, Wallace JC, Hood L
99380589
COMMENT University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.hnsc.washington.edu
Plate: 2082 row: B column: 9
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 592.

```

```

FEATURES
Source
Location/Qualifiers
1..592
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Plate=2082 Col=9 Row=B"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 193 a 115 c 146 g 125 t 13 others
ORIGIN

```

```

Query Match 21.6%, Score 205; DB 17; Length 592;
Best Local Similarity 61.2%, Pred. No. 4,le-42;
Matches 325; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

```

```

QY 175 CCCATGTTCTTCTCCAGCAGCTCCCATGAGCAGCATGTTGCTGTACATT 234
Db 592 CGCATGTTCTTCTCCAGCAGCTCCCATGAGCAGCATGTTGCTGTACATT 533

```

```

QY 235 GTGCCAAGATGGCAGCAACTTCTGTGCGAGGAAATCCATTCCTTTGCGCTGT 294
Db 532 GTACCCAGATGGCTTCAACTCTGTGTGCGAGCAAGTTCATTTCTATGCGTGTGT 473
QY 295 GGCATACAAATTTGGCTTTTGTCTCTTGTGGATCTGAGGAGGCTTGTGCGACTG 354
Db 472 GCCACACAAATTTCTGTGTATCACTCTTGGCTGTAAATGCTTTCTGTGTGCTGT 413
QY 355 ATGGCTTATGACACATCATGTTGCGTGTAGGACCCACCTTACATATCCATCTCATGAAT 414
Db 412 ATGCTTATGACCTCATATTCATTTGCCATTTGCCACCTGTAGATATACCAATTCATGAGA 353
QY 415 CAGAGGCTGTCTCCAGATTTACTGAGAGCTCTCGGAGCTTTGGATATAGATAGAGTG 474
Db 352 CCCAATTTTGTACATCATATGACCTCTCTGATGTTGGCTCTACAGATGANATC 293
QY 475 ATTCAGATGGTGGCAGCATGAGCTTACTTACTGTGCTGCTCAAGAGAGCTGATCATTTT 534
Db 292 ATTATGCTGAGCCACATTTTCTCTCTCTACTGTGTGGTGTGGAAATAGCCACTTC 233
QY 535 TTCTGTAGGTACAAAGCTTTATTTGAGCTGCGCTGTGCAAGACACTTCCCTTTTGACAC 594
Db 232 TTCTGTAGGTACCTTCCCACTAATCTCTCATGATGATACATCATATCTGAAGAAG 173
QY 595 CTCTCTTTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 654
Db 172 GTTATTTTCAATGCTCTGTATAGTATGCTTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 113
QY 655 TATGCTTGCATCTTACGAGGCTGTGCTCCGATACGCTGTCTGTCTGTCTGTCTGTCTGT 705
Db 112 TATCAGAGATTAATTTGCGCTGTATTCATCATGATGATCTGTGAGAGGCTCG 62

```

```

RESULT 9
A0425578/c 448 bp DNA linear GSS 24-MAR-1999
LOCUS CITBI-E1-2567E19.TR CITBI-E1 Homo sapiens genomic clone 2567E19,
DEFINITION DNA sequence.
ACCESSION A0425578
VERSION A0425578.1 GI:4495666
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuza,H., Simon,M. and
Venter,J.C.
Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
Unpublished (1997)
Other-GSSs: CITBI-E1-2567E19.TR
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetlgr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

```

```

FEATURES
Source
Location/Qualifiers
1..448
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="2567E19"
/clone_lib="CITBI-E1"
/sex="male"
/note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;

```

```


```


| | CalTech Human BAC Library D* | | | |
|------------|------------------------------|---|----|--------------|
| BASE COUNT | 149 | a | 80 | c 126 g 93 t |
| ORIGIN | | | | |

| | | | | |
|---------------------------|-------|-------------------|----------|------------|
| Query Match | 21.2% | Score 200.6 | DB 17 | Length 448 |
| Best Local Similarity | 66.1% | Pred. No. 4.9e-41 | | |
| Matches 290, Conservative | 0 | Mismatches 149, | Indels 0 | Gaps 0 |

21.28; Score 200.6; DB 17; Length 448;
66.19; Prof No 4 08-41

Best Local Similarity 00.16; Prev. NO. 4.9E-41

Matches 290; conservative 0; mismatches 149; indels 0; gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 104 | TCTTCACAGTAGGCCCTCTGTGGGAAATGTCCTCCATCTTCCTCATCTCTGAGAGCGTG | 163 |
| Db | 444 | TCTTTTGTGTGGCATTCAGGGAACCTGTGATGTGTTCTCCTCATATACCTGGACACCC | 385 |
| QY | 164 | GACTTCACACCCCACTGACTCTCTCTCCAGCAGCTCTCCCTCATGACCTCATGTTGG | 223 |
| Db | 384 | AGCTCCACACCCCACTGTACTCTCTCCAGCAACTGTCCCTCATGACCTCATGTCTCA | 325 |
| QY | 224 | TCTGTAACTATGTGGCAAAATGGCAGCAACTCTCTGTGTGGCAGGAAGTCCATCTCT | 283 |
| Db | 324 | TCTGACCAACCGTACCCAGAGTGGCTTCAACTACTCTGTGTGGCAAGTCCATTTTCA | 265 |
| QY | 284 | TTTGGGCGCTGTGGCATACAAATTTGGCTTTTGTCTCTCTGTGGAGTCTGAGGGGCTCT | 343 |
| Db | 284 | TGGCGTGTGTGTACACAAATTTCTTCTATATATCACTGTCTGGCTGTGAATGTTTC | 205 |
| QY | 344 | TGCTGGGACATAGGCTTATGACCACTACGTGTGGCGCTTAGCCACCCACTTCACTATCCA | 403 |
| Db | 204 | TTTTGGCTGTATGGCTTATGACCGCTATTTGGCTATTTGGCAACCCCTCATAGATATACCA | 145 |
| QY | 404 | TCTCATGATATACAGAGGGTCTGTCTCCAGATTACTAGGAGCTCTGGGCTTTGGAGTAA | 463 |
| Db | 144 | ATCTCATATAAATCTTAAATTTGTGTGACTTATGTCTTCTCTGTGATCTCTGGGCTCTTA | 85 |
| QY | 464 | TAGATGAGTATTCAGATGTGTGGCAGCCATGAGGCTTACCTTACCTTATGTGACTCAAGAGCG | 523 |
| Db | 84 | CAGATGAGTATTCATGTGCTGTAGCCACATTTTCTCTTCTCTTTGTGGGTCTCGGANA | 25 |
| QY | 524 | TGATCACTTTTCTGTGA | 542 |
| Db | 24 | TAGCCACTTCTCTGTGA | 6 |

| | |
|------------|---|
| RESULT | 10 |
| B0888002 | |
| LOCUS | |
| DEFINITION | B0888002 915 bp mRNA linear EST_16-AUG-2007 ACNCCOURT_8742018 NIH_MGC_129 Mus musculus cDNA clone IMAGE:639037 |
| ACCESION | B0888002 B0888001 |
| VERSION | B0888002.1 GI:22280016 |
| KEYWORDS | EST. |
| SOURCE | house mouse. |

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1-915 |

```

/clone_lib="NIH_MGC_129"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: olfactory epithelium; Vector:
pCMV-Sport6.1 codb; Site_1: EcoRV;
unidirectionally. Primer: Oligo dt. Average insert size
2.2 Kb. Constructed by Resgen, Invitrogen Corp. Note: This
is a NIH_MGC library."
BASE COUNT
ORIGIN
181 a      287 c      201 g      246 t

```

| | | | | |
|------------------------|--------------|------------|-----------------|-------------|
| Query Match: | 20.9% | Score 198: | DB 14: | Length 915; |
| Best Local Similarity: | 53.7% | Pred. No. | 3.3e-40: | |
| Matches 433; | Conservative | 0; | Mismatches 370; | Indels 3; |
| | | | | Gaps 1; |

| | | | |
|--------|------------|--------|------------|
| 20.98; | Score 198; | DB 14; | Length 915 |
| 53.78; | Prod no 3 | 30-40; | |

Best Local Similarity 33.78; Pred. NO. 3.3E+40;
 Notebook 433; Construction 0; Wavelength 370; T-3-1-3; 3-1-1

malchies 433; conservative 0; mismatches 3/0; indels 3; gaps 1;

| | | | |
|----|-----|---|-----|
| QY | 124 | GGGAAATGCTCCATCTCATCTTCCTCATCTTACCTGGAGAGCTGGAGCTTCAACACCCCATGTAC | 183 |
| Db | 4 | GGAAACGCTCATTTATGCTGACAAATGATGCTCCAGCCCAACCCCTGCATPACCCCTATGTAC | 63 |
| OY | 184 | TTTCTTCTCAGCACACTCTCCCTCATGAGACCTCATGTTGGTCTGTAACTATGTCCAAAG | 243 |
| Db | 64 | TTTCTTCTCGTCAACCTCTGGCTGTGGTGGACATCTCTTGGACCTCCACACTCTCAACCCAG | 123 |
| OY | 244 | ATGGCAGCCAACTTCCTCTCTGGGAGGAAAGTCAATTCCTTGTGGGCTGTGGGATCAAA | 303 |
| Db | 124 | CTACTGGACACCAATGCTGGGGGAGAGA--CCATCTCTTACGGGGGCTGCATGGCCAG | 180 |
| OY | 304 | ATTGGCTTTTTTGTCTCTCTTGTGGGATCTGAGGGGCTCTTGTGCGGAGACTCATGGCTTAT | 363 |
| Db | 181 | CTCTTCTTCTTCAATGATGTCACTGGGGGTGAGGCTTCTACTCTTCTCAGCATGCGCTAT | 244 |
| OY | 364 | GACCACATACGTGGCGGCTTATGACCAACCACCTTACTATCCATCTCATGATCAGAGGTC | 423 |
| Db | 241 | GACCGCTTTGGGCACTCTGCTGGCCCTCTGCTATTATGTAACCTGGAATGGGCCCCAGATG | 300 |
| OY | 424 | TGCTCCAGATTACTGGGAGAGCTCCGCGGCGCTTGGGAAATAGATGAGATGCATAG | 483 |
| Db | 301 | TGTGCAATTTCTGGCTGGCATTGTCTGGGCTATCAGCCTGATTAACACCAAGCATTAACCTCA | 360 |
| OY | 484 | GTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGGTGATCACTTTTCTGTGAG | 543 |
| Db | 361 | AGCGTGGTACGTCATCACTTCCTGGAGCTCAATGTGGTGTAGAGCACTTCTGTGAG | 420 |
| OY | 544 | GTCACAGCTTTATTTAAGAGCTGGCTGTGACAGACACTTCCCTTTTGTACACCCCTCTCTTT | 603 |
| Db | 421 | ATTCCCCCACTGTTGAACCTCTCTGTGGCTCAACACAAATTTGAACGAGGCGCATGGCTTTC | 480 |
| OY | 604 | GCTGCTGTGCTTCATGACCTCTCTCTTCCCTTCCATCATCATGAGCGCTCATATGCTTGC | 663 |
| Db | 481 | GCTGCATATGTGTCTGTGCTGTGAGGACACTTCTGTGATCATCTCTCTATGCTTT | 540 |
| OY | 664 | ATCTAGAGGGCTGTGCTCCGAATACGCTCTGCTCAGGCTGGAAAAAAGCCCTGGCCACC | 723 |
| Db | 541 | ATTGTGGCCACATCTGTAAGATCCGCTCAAGCTGAGGGAAGCAAGCGCTTCCACC | 600 |
| OY | 724 | TGCTCTCCCAACCTAACAAGCTGTACCCCTCTTCTATATGGGAGCGCATGTTTCACTGTACTG | 783 |
| Db | 601 | TGCTGTGACACCTCATCTGTGCTGTACCATGTACATCTCACTGTCACTCATCACTT | 660 |
| OY | 784 | AGGCGCTAGGCGCTACCGGGCCCTTAGCCATATACAAGGTGGCTCTTATCTTCAACAGTC | 843 |
| Db | 661 | CGCCCTTTCATCCACTACTCACTCAACAAGACAAAGGTGGTGTCTCATCATCTACACCTCA | 720 |
| OY | 844 | CTTACTCCCATGCTGAACCCCTCATTTACAGCTTGAAGGATGGGAGGTGATGGGGCA | 903 |
| Db | 721 | GTGGCACCCACCTTGAACCCCTCATCTACACTGTGAGGAATAAGATGTCCAAAGTTGCA | 780 |
| OY | 904 | CTGAGGAAGGGGCTGAGACCGCTGAG | 929 |
| Db | 781 | CTCCGGAGACTTTTCTCTCTGTGCTG | 806 |

RESULT 1
BI912805

| | | | |
|----|-----|---|-----|
| Db | 181 | TCCTGGCGTTATGGCAATATGACCGCTAACGGGGGCTGTCTGCAAGCCCCCTGGACTACACCG | 240 |
| Qy | 404 | TCCTCATGAAATCACAGAGGCTCTGTCTCCAGATTACTGGGAGCTCCTGGGCCCTTTGGGATAA | 465 |
| Db | 241 | TCATCATATGCAATCCGTCGTCTGCGGGGACGGGCTTCAATGGCATGGTTGATGCGCGCTTTG | 300 |
| Qy | 464 | TAGATGGAGTATTTCAGATGGGTGGCAGCATGGGCTTACCTTACTGTGGGTCACAGGAGCG | 522 |
| Db | 301 | GCAATTCTCTATATATGAGCACCCACAGAGTTGATGTTGGCCCCGCTGTGGACACAGGCGGG | 360 |
| Qy | 524 | TGGAATCACTTTTCTGTGAGGTATACAAGCTTTATGAAGCTGGCTGTGCAGACACTTCCC | 583 |
| Db | 361 | TGGACAACTTCTCTGTGAGATGGCGGGCCGATTGGCATATGCGCTGTGTGTGACACTATGG | 422 |
| Qy | 584 | TTTTTGAACCCCTCCTTTTGCTGTGTGTCTTCATGTCTTCCTTCCCTTCATCA | 643 |
| Db | 421 | CCTTGGGAAGCACTGGCTTTTTCGTTGGCGATTTTATCATCTTGGCACCGGCTCATCTCCA | 480 |
| Qy | 644 | TCATGGCGCTCTATGCTTGGCATCTCTAGGGGCTGTGTCTCCGAATACGCTGTGCTTAGGCT | 703 |
| Db | 481 | TCTTATATCTCTATAGGTATCAATTCGCCGAGAGTGTAGGATCAAGTCAGCTCTCGGGC | 540 |
| Qy | 704 | GGAAATGAAGCCCTGGCAACCTGCTCTCCCACTTAACAGGCTGACCTCTTATATGAGGG | 765 |
| Db | 541 | GAGCGAAAGCCTTCAACACTTGCACGCTCCCACTTAATATCGTTGTGTCTCTCTCTTCTACGGTA | 600 |

BASE COUNT
ORIGIN

116 a 189 c 139 g 174 t

```
/clone="umcGC2M0180J22"  
/clone_lib="mouse 10kb plasmid umcGCM library"  
/sex="Male"  
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptored DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of PWD42 (g114732114|9b|AF129072.1), a copy-number  
inducible derivative of plasmid RL. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptored mouse DNA was annealed to  
adaptored vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."
```

| | | | |
|-----------------|---|--|-------------------------------------|
| Db | 601 | CCATCATATATACATGACTCCAGCGCGCAAACTTAACTCCCGACAGCACGAGGCGCAAGTTT | 660 |
| QY | 824 | CCCTCATCTCTACACA | 840 |
| Db | 661 | CTCACTCCTTTTCTTACA | 677 |
| RESULT 13 | | | |
| LOCUS | A2868627 | | |
| DEFINITION | AZ868627 | 618 bp | DNA linear |
| ACCESSION | 2M0180J22F | Mouse 10kb plasmid | U08C1M library Mus musculus genomic |
| VERSION | A2868627 | | |
| KEYWORDS | A2868627.1 | GI:13072131 | |
| SOURCE | GSS. | | |
| ORGANISM | house mouse. | | |
| REFERENCE | Mus musculus | | |
| AUTHORS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 618) | | |
| TITLE | Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R. | | |
| JOURNAL COMMENT | Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts | | |
| | Unpublished (2000) | | |
| | Contact: Robert B. Weiss | | |
| | University of Utah Genome Center | | |
| | Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA | | |
| | Tel: 801 585 5606 | | |
| | Fax: 801 585 7177 | | |
| | Email: ddunn@genetics.utah.edu | | |
| | Insert length: 10000 | | |
| | Std Error: 0.00 | | |
| | Plate: 0180 | | |
| | row: J | | |
| | column: 22 | | |
| | Seq primer: CGTGTAAACGACGCGGCAGT | | |
| | Class: plasmid ends | | |
| | High quality sequence stop: 618. | | |
| | Location/Qualifiers | | |
| | 1. 618 | | |
| FEATURES | | | |
| source | /organism="Mus musculus" | | |
| | /strain="C57BL/6J" | | |
| | /db_xref="taxon:10090" | | |

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2003, 03:35:05 ; Search time 73 Seconds
(without alignments)
3982.597 Million cell updates/sec

Title: US-09-634-109D-1
Perfect score: 948
Sequence: 1 atgggaagatgggtgaacca.....ggattgcagcagcactga 948

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724.

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|------------------|
| 1 | 388.8 | 41.0 | 1290 | 2 | US-08-827-291A-1 |
| 2 | 245 | 25.8 | 1713 | 2 | US-08-467-948A-1 |
| 3 | 245 | 25.8 | 1713 | 2 | US-08-467-947A-1 |
| 4 | 235.2 | 24.6 | 966 | 3 | US-08-748-506-7 |
| 5 | 233.6 | 24.6 | 966 | 3 | US-08-748-506-8 |
| 6 | 233.2 | 24.6 | 966 | 3 | US-08-748-506-5 |
| 7 | 227.4 | 24.0 | 966 | 3 | US-08-748-506-6 |
| 8 | 208 | 21.9 | 1080 | 4 | US-09-668-680-9 |
| 9 | 203.4 | 21.7 | 900 | 4 | US-09-085-371-5 |
| 10 | 203.4 | 21.5 | 1297 | 4 | US-09-668-680-11 |
| 11 | 201.4 | 21.2 | 1062 | 4 | US-09-668-680-10 |
| 12 | 186.2 | 19.6 | 984 | 3 | US-08-748-506-9 |
| 13 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 14 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 15 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 16 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 17 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 18 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 19 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 20 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 21 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 22 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 23 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 24 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 25 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 26 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |
| 27 | 182.4 | 12.9 | 963 | 4 | US-09-668-680-13 |

| | | | | | | |
|----|------|-----|------|---|--------------------|---------------------|
| 28 | 56.8 | 6.0 | 3211 | 4 | US-09-357-014-8 | Sequence 8, Appl1 |
| 29 | 56.8 | 6.0 | 3901 | 2 | US-08-574-959A-6 | Sequence 6, Appl1 |
| 30 | 56.8 | 6.0 | 3901 | 4 | US-09-357-014-6 | Sequence 6, Appl1 |
| 31 | 47.8 | 5.0 | 251 | 4 | US-09-605-785-352 | Sequence 352, Appl1 |
| 32 | 47.8 | 5.0 | 251 | 4 | US-09-439-313-352 | Sequence 352, App |
| 33 | 47.8 | 5.0 | 251 | 4 | US-09-352-616A-352 | Sequence 352, App |
| 34 | 47.6 | 5.0 | 1270 | 4 | US-08-387-805-1 | Sequence 1, Appl1 |
| 35 | 46.6 | 4.9 | 921 | 1 | US-08-722-001-17 | Sequence 17, Appl1 |
| 36 | 46.6 | 4.9 | 1567 | 1 | US-08-722-001-24 | Sequence 24, Appl1 |
| 37 | 46.6 | 4.9 | 1738 | 1 | US-08-334-698-3 | Sequence 3, Appl1 |
| 38 | 46.6 | 4.9 | 1738 | 1 | US-08-228-932-3 | Sequence 3, Appl1 |
| 39 | 46.6 | 4.9 | 1738 | 1 | US-08-468-939-3 | Sequence 3, Appl1 |
| 40 | 46.6 | 4.9 | 1738 | 2 | US-08-406-855A-3 | Sequence 3, Appl1 |
| 41 | 46.6 | 4.9 | 1738 | 2 | US-08-722-190-3 | Sequence 3, Appl1 |
| 42 | 46.6 | 4.9 | 1738 | 3 | US-08-244-354-3 | Sequence 3, Appl1 |
| 43 | 46.6 | 4.9 | 1738 | 3 | US-09-206-899-3 | Sequence 3, Appl1 |
| 44 | 46.6 | 4.9 | 1738 | 4 | US-09-444-783-3 | Sequence 3, Appl1 |
| 45 | 46.6 | 4.9 | 1738 | 4 | US-09-688-415-3 | Sequence 3, Appl1 |

ALIGNMENTS

RESULT 1
US-08-827-291A-1
; Sequence 1, Application us/08827291A
; Patent No. 5874243
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Sathe, Ganesh
; TITLE OF INVENTION: NOVEL OLRC15 RECEPTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: PA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,291A
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T.
; REGISTRATION NUMBER: 30,954
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5015
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1290 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-827-291A-1
; Query Match 41.0%; Score 388.8; DB 2; Length 1290;
; Best Local Similarity 64.4%; Pred. No. 7.8e-103;
; Matches 582; Conservative 0; Mismatches 322; Indels 0; Gaps 0;
; Oy 9 ATGGTGACACAGTCCACAGATGGCTTCTCTTGGGCACTTTTCCACAGCA 68

| | | | |
|----|-----|---|-----|
| Db | 518 | TATTTCATCATCATGACCTGGAAAGTCTGCATCACTCTGGGCACTCTTCGAGATGT | 577 |
| QY | 457 | GGGATTAATAGATGAGATGATTCAGATGTTGGCAGCCATGGGCTTACTTACTGTGCTCA | 516 |
| Db | 578 | GGCTCCCTCCCTGGCTAATGGTCCATAGTAGGCTCATCCTAAGACTGCCCCCTTTTGGGCC | 637 |
| QY | 517 | AGGAGCGTGCATCACTTTTCTGGAGGTACAAAGCTTATTGAAGCTGGCCTGGCAGAC | 576 |
| Db | 638 | CGTGAATACCAACCACTTCTTGTGGAATCTGTGCTCTCCTCAGGCGTGGCCTGTGAT | 697 |
| QY | 577 | ACTTCCCTTTTGAACACCCTCTCTTTCCTGCTGTGCTGTCTTCACTCTCTCCCTTC | 636 |
| Db | 698 | ACCGGCTCAACCAAGGTGTGCTATCTTTGAAGCCGCAATGTTCACTCCGGGAGCACCTC | 757 |
| QY | 637 | TCCATCATCATGAGGCTCCTATGCTTGCATCCTTAGGGGCTGTGCTCCGAATAGCCTGCT | 696 |
| Db | 758 | TGCGTGTGCTGCTGCTCTCTACTCAACATCTTGGGGGACATCCTGAGAGATCCAGTCTGGG | 817 |
| QY | 697 | CAGGCGTGAAGAAAAAGCCCTGGCCACCTGCTCTCCACCACTAAGCTGTACACCTTTC | 756 |
| Db | 818 | GAGGAGCGCAGAAAGGCTTCTTCACCTGCTCTCCACCTCTGCGTAGAGTGGAGCTCTTC | 877 |
| QY | 757 | TATGGGAGCAGCATGTCATGTACTTACCTAGAGCCTAGGCGCTACCGGGGCCCTTACCATGAC | 816 |
| Db | 878 | TTTGGSAAGCCATGTCATGATGATGATGAGCCCTTAAGTCCGCAATCCTGAGAGACACAG | 937 |
| QY | 817 | AAGGTGGCCTTATCTTCTACACAGTCTTACTTCCCATGCTG-----AACCCCTCATTTA | 872 |
| Db | 938 | AAGGTCCCTTTTCTTATTTTTACATGTTCTTTCAACCCCATGCTTAACCCCTGATTTA | 997 |
| QY | 873 | C-AGCTTGAGGATGGGAGG 892 | |
| Db | 998 | CAACCTTGAGGATGTAGAGG 1018 | |

RESULT 3
 US-08-467-947A-1
 Sequence 1, Application US/08467947A
 Patent No. 6090575
 GENERAL INFORMATION:
 APPLICANT: LI, YI
 APPLICANT: CAO, LIANG
 APPLICANT: NI, JIAN
 APPLICANT: GENTZ, REINER
 APPLICANT: BULL, CAROL J.
 APPLICANT: SUTTON III, GRANGER G.
 APPLICANT: ROSEN, CRAIG A.
 TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
 TITLE OF INVENTION: Coupled Receptor GPR1
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 STREET: 1100 NEW YORK AVE., NW, SUITE 600
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,947A
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/04079
 FILING DATE: 30-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: STEFFE, ERIC K.
 REGISTRATION NUMBER: 36,668

| | | |
|----|--|--|
| | REFERENCE/DOCKET NUMBER: | 1488..1140002/EKS/KLM |
| | TELECOMMUNICATION INFORMATION: | |
| | TELEPHONE: | 202-371-2600 |
| | TELEFAX: | 202-371-2540 |
| | INFORMATION FOR SEQ ID NO: 1: | |
| | SEQUENCE CHARACTERISTICS: | |
| | LENGTH: | 1713 base pairs |
| | TYPE: | nucleic acid |
| | STRANDEDNESS: | both |
| | TOPOLOGY: | both |
| | MOLECULE TYPE: | cDNA |
| | FEATURE: | |
| | NAME/KEY: | CDS |
| | LOCATION: | 116..1003 |
| | US-08-467-947A-1 | |
| | Query Match | 25.8%; Score 245; DB 3; Length 1713; |
| | Best Local Similarity | 56.8%; Pred. No.3.2e-61; |
| | Matches 489; Conservative 1; Mismatches 366; Indels 5; Gaps 2; | |
| Oy | 37 | TTCCTCCCTGTTGGGCATCTTTCACACAGCCAGACTGACCCTGTCTCCTCATCTG |
| Db | 158 | TTCCTCCCTACACCGGATTTCCTCTGGGCCCAAGATTACAGATCTCTCTTGGGCTTC |
| Oy | 97 | ATGTGTGCTTCACAGTGACCCTCTGTGGGAATGTCTCTCATCTTCTCATATCACTG |
| Db | 218 | TCCCCTGTTCTATGTCTTACACCTGCTGGGGAAATGGAGCACCATCTGGGGCTCATCTCAGT |
| Oy | 157 | GACGCTGGACCTTCACACCCCCATGTACTTCTCCACAGCAGCTGCMCCCTCATGGACCTC |
| Db | 278 | GACTCCAGACTTCACACCCCCATGTACTTCTCTCTCTACACCTGGCCGTGTCACATC |
| Oy | 217 | ATGTTGCTGTGAACATTGTGCGCAAAGATGGCAGCCAACTTCCTCTGTGGCAGAATGC |
| Db | 338 | GCCATGCTGCACACAGTAGTGGCCAGATGCTGTGAACCTCTCATCCACACCAAGCC |
| Oy | 277 | ATCTCTTTGTGGGCTGTGGCATTAACAATTGGCTTTTTGTCTCTCTGTGGGATCTGAG |
| Db | 398 | ATCTCTCTTTGGCTGTGACATGACACTTACCTTCTCTTTTAACTTTGGTTCACATACTGAA |
| Oy | 337 | GGGCTCTTGCTGGGAGCTATGCTTATGACCACTACGTGCGGCTTAAGCACCACCTTAC |
| Db | 458 | TGCCCTCGTTGGTGTCTGATGTCTACAGATGGGTACGTGGCCATCTGCCACCTCTCGA |
| Oy | 397 | TATCCCATCTCATGAATCAGAGGCTCTGTCTCCAGATTACTGGAGACTCTGGGCTTT |
| Db | 518 | TATTTTCATCATCATGACCTGGAAAGTCTGCATCACTCTGGGGCATCACTTCGGAGCATGT |
| Oy | 457 | GGGATAATAGATGAGCTGATTCAGATGGTGGGAGCCATAGGGCTTACCTTACGTGGCTCA |
| Db | 578 | GGCTCCCTCTGGCTATATGATCATGTGAGCTCAATCTTAAGCTCCCTTTGTGTGGCT |
| Oy | 517 | AGGAGCGTGGATCACTTTTCTGTGAGGTACAAAGCTTATTGAAAGCTGGCCGTGTACAG |
| Db | 638 | CGTGAATCAACCACTTCTGTGTGAANAATCGTCTGTCCTCAGGCTGGCCGTGCTGAT |
| Oy | 577 | ACTTCCCTTTTGTACACCCCTCTCTTTGCTGTCTGTCTGTATGCTCTCTCTCCCTTC |
| Db | 698 | ACCTGGCTCAACCAAGTGGTATCTTTTGAAGCTCTCATATGTTATCTCTGGTGGACACAC |
| Oy | 637 | TCCATCATCATGAGCTCCATATGCTTGCATCTCAGAGGGGCTGTCTCCGAATAGCTCTGCT |
| Db | 758 | TCCCTGTGTGTGTCTCTCTACACACATCTCTGGGGGGGATCTCTGAGATCCAGTCTGG |
| Oy | 697 | CAGGCTGTGAAAAAGCCCTGGCCACTGTCTCTCCACACTCAACAGCTGTACCTCTTC |
| Db | 818 | GAGGGCCCGAGAAAGCCCTTCCACCTGCTCTCTCCACCTCTGTGTGTGGAGACTCTTC |
| Oy | 757 | TATGGGGGAGCATTTGATGACTGTAGGGCTTAGGGCTAACCGGGCCOCTAGCATGAC |
| Db | 878 | TTTGGSAGCGCATGTGATGTACATGAGCCCTTAAGTCCCGGCATCTGTGAGAGACGACAG |
| Oy | 817 | AAGGAGGCTCTATCTTACACATCTCTTACCTCCATCTG---AACCCCTCATTTTA |

DB 938 AAGGTCCTTTCTATTATTACAGTTCCTTCACCCCGATGCTTAACCCCTGATTTTA 997
OY 873 C-AGCTTGAGAGTGGGAGG 892
DB 998 CAACCTGAGAGATGTAGAG 1018

RESULT 4
US-08-748-506-7
Sequence 7, Application US/08748506
Patent No. 6159707
GENERAL INFORMATION:
APPLICANT: Ronnett et al.
TITLE OF INVENTION: NOVEL SPERM RECEPTORS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leydig, Voit & Mayer, Ltd.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: IL
COUNTRY: US
ZIP: 60601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/748,506
FILING DATE: 08-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,751
FILING DATE: 09-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 74940
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-616-5600
TELEFAX: 312-616-5700
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-748-506-7

Query Match 24.8%; Score 235.2; DB 3; Length 966;
Best Local Similarity 54.8%; Pred. No. 1.7e-58;
Matches 465; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

OY 75 CCTTGTCCCTCTCTGACGATATGCTGCTTCACAGTGGCCCTCTGTGGAAATGCT 134
DB 90 CTTCTCTGCTTACCCCTATCTCTCATGTGTTAGTATCATCAACGAAATACTCT 149
OY 135 CCTCATCTTCTCATCTACCTGAGCTGAGCTTCACACCCCGATGCTTCTCTGAG 194
DB 150 CATACCCCTGCTATTGTTACACAGTCACTCTACACACCCCGATGCTTCTCTGAG 209
OY 195 CCAGCTCTCCCTCATGAGACCTGATGCTGTGCTGTACATTTGTGCAAAAGTGGACCAA 254
DB 210 CAACCTGCTCTCTGAGATGCTATGCTGCTGTCTCATACCAAGATGCTGACAG 269
OY 255 CTTCTGTCTGACGAGAGTCACTCTCTTGTGGCTGTGACATACAAATGCTTTT 314
DB 270 CTTGTGAGTGGCCCGAGGAGATCTTTGGAGGCTGTGCTTCAAGATGCTTCTT 329
OY 315 TGTCTCTTGTGGAGATGAGAGGCTCTTGTGGAGCTCATGCTTATGACACTAGCT 374
DB 330 CATATCTTTGTATACGATGCTGCTATTTGGCAGCATGCGCTTTGACCGCTATAT 389

OY 375 GGCGTTAGCCACCACTTCACTATCCATCCATGANTGANTGAGAGGCTGCTCCAGAT 434
DB 390 GGCTATATGTTCCCACTCACTCATATGACACCGGANTGATGCTGGGATATGCTATTT 449
OY 435 TACTGGAGCTCTGGGCTTTGGATATATGATGAGTGTATGATGATGAGTGGACCAT 494
DB 450 GGCAATTTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 509
OY 495 GGGCTTACTTACTGAGGCTCAAGAGAGTGTATGATGATGATGATGATGATGATGAT 554
DB 510 CTTCTTGAATCTTGTGACCTGTGAGATGATGATGATGATGATGATGATGATGAT 569
OY 555 ATTGAAGCTGGCTGTGACACCTTCTCTTTTGAACACCTCTCTCTGCTGTGT 614
DB 570 CTTGGACTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 629
OY 615 CTTGATGCTTCT 674
DB 630 GCTTGCATATTTAGTCCATTTTACTGATTCATTTCTTCTTCTTCTTCTTCTTCT 689
OY 675 TGTGCTCCGATAGCTCTGCTCAGGCGCTGGAAGAGCCCTGGCAGCTGCTCTCCA 734
DB 690 AGTGTGTGATGCTCTTCACTGAGGCGGCGCATTAAGCTCTCTTACTCTGTCATCTCA 749
OY 735 CTTACAGCTGTACACCTCTTCTATGAGGCGACCATGTTGATGATGATGATGATGAT 794
DB 750 CTTACTTGTAGTACACTTCTTATGCTCAGACATCTGCACTTGTGAGTGTCAAGTC 809
OY 795 CTACGGGGCCCTTACCATGACAGAGTGGCTCTTCTTCTTCTTCTTCTTCTTCT 854
DB 810 TAGCCACTCAGCAGAGAGTGGACAACTTGGCCCTCTTCTTCTTCTTCTTCTTCT 869
OY 855 GCTGAACCCCTCATTTTACAGCTTGAAGAGATGAGAGTGGAGGAGTGGAGTGGAG 914
DB 870 GCTGAATCCCATCTTACAGCTTGAAGAGATGAGAGTGGAGTGGAGTGGAGTGGAG 929
OY 915 GCTGAGC 922
DB 930 TCTGGCC 937

RESULT 5
US-08-748-506-8
Sequence 8, Application US/08748506
Patent No. 6159707
GENERAL INFORMATION:
APPLICANT: Ronnett et al.
TITLE OF INVENTION: NOVEL SPERM RECEPTORS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leydig, Voit & Mayer, Ltd.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: IL
COUNTRY: US
ZIP: 60601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/748,506
FILING DATE: 08-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,751
FILING DATE: 09-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 74940
TELECOMMUNICATION INFORMATION:

| | | | |
|----|-----|---|-----|
| QY | 435 | TATGGGAGGCTCTGGGGCTTTGGATTAAGATGAGTATTCAGTATGGTGGAGGCAT | 494 |
| Db | 450 | GGCAATTGTTTCATGGGGAAATGGATGCATAGTAAAGCTGGGCAAAACCAATTTATTTT | 509 |
| QY | 495 | GGGCTTACCTTACATGCTGGCTTCAGAGGCGTGGATCACTTTTCTGTGAGGTACAAAGCTTT | 554 |
| Db | 510 | CTCTGTGAACCTTCTGTGGACCTGTGAATAAGACCACTTCTCTGTGACCTTCCACCTCT | 569 |
| QY | 555 | ATTGAGGTGGCGCTGGGAGACACTTCCCTTTTGGACACCCCTCTCTTGTGCTGTGTG | 614 |
| Db | 570 | CTTGGCACTTCTCTGTGGATATACATCCCAAAAGAGGCTGCCATCTTGTGTGAGCAGT | 629 |
| QY | 615 | CTTCATGCTTCCTCTCCCTTCTCCATCATCATGCGCCTCTATGCTTGCATCCTTAGGGG | 674 |
| Db | 630 | CCCTGCATATCTAGCCCAATTTTGGTGCATCATTTATTCTTATGTCAAAATTCATATGC | 689 |
| QY | 675 | TGTGCTCGGAATPAGCGCTGTGCTGAGGCGCTGGAAAAAGCCCTGGCCACCTGCCTCCCA | 734 |
| Db | 690 | AGTGTCTMGATGCTTCTTACCTGAGGGGCGCATAAAGCTCTTTCACCTGTGTGCTCA | 749 |
| QY | 735 | CCCTAACGCTGTCAACCTCTTCTCATATGGGGGAGCCCATTTATGACTGAGGCGCTAGGCG | 794 |
| Db | 750 | CCCTACTGTATGACACTTTTATTTTGTGCTCAGCATGTATTACCTATTTTGAAGGCCAAGTC | 809 |
| QY | 795 | CTACCGGGCCCCCTAGGCATAGACAAGGTGGCCTCTATCTTTCTAACACAGTCTTACTCCAT | 854 |
| Db | 810 | TAGCACTACACAGGAATGACCAATTTCTTGGCCCTCTTTACACACATTAATGCATTCAT | 869 |
| QY | 855 | GCTTAACCCCTCATTTACAGCTTGAAGATGGGAGTATGGGCGCTGAGAGAAGG | 914 |
| Db | 870 | GCTGAACCCCTATCATATATAGTTTAAGAACAAAGGAAGTCAAGGACACACATGAGAGAAGAC | 929 |
| QY | 915 | GCTGGAGC 922 | |
| Db | 930 | TCTGGGCC 937 | |

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 210121.427C9
 ; CURRENT APPLICATION NUMBER: US/09/439,313
 ; CURRENT FILING DATE: 1999-11-12
 ; NUMBER OF SEQ. ID NOS: 575
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ. ID NO 526
 ; LENGTH: 963
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-439-313-526

Query Match 12.9%; Score 122.4; DB 4; Length 963;
 Best Local Similarity 49.0%; Pred. No. 6.7e-26;

Matches 327; Conservative 0; Mismatches 341; Indels 0; Gaps 0;

```

QY 111 AGTGGCCCTGTGGGAATGTCCTCCATCTTCCATCTACTGAGCGCTGACTTCA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 105 AGTGGCAATGTTGGAACTGATCGTGTCTTCATGTAAGAGGAGGACGACCTGCA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 171 CACCCCATGTAATCTTCTCCACGACCTCCCTCATGAGCTCATGTTGTTGTTAA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 165 CGCTCCATGTAATCTTCTCTGATGCTTGCAGCATGACCTGCTTATCCACATC
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 231 CATTGTCGCAAGAGGACCACTCTCTGTCGAGAGAGTCCATCTCTTGTGGG
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 225 CACCATGCTTAAGATCTGCTTCTGTTGATTCCGAGAGATTAAGCTTGAAGG
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 291 CTGCGCATCAATATGGCTTTTGTCTCTCTGTGGATCTGAGGCGCTTGTGGG
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 285 CTGCTTACCAAGATGTTCTTATTCATGCTCCCTCCAGCCATTGAATCCACATCTGCT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 351 ACTGATGCTTATGACCACTAGTGGCGCTTACGACCACTTACTATCCATCTCAT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 345 GGCATGCTCTTACCGCTTATGTCGATCTGCACTGCGCATGCTGAGTCT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 411 GAATCAGAGGCTGTCTCCAGATTACTGGAGCTCCTGGCCCTTGGGATTAATGATG
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 405 CAACATATACAGTACAGCCAGATTGCGATGCTGTGCTCGCGGATGCCCTTTT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 471 AGTATTCAGATGCTGACGACCATGAGCTTACTTACTGTGCTCAAGAGCGTGATCA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 465 TTTTCCACCTCTGCTGATCAAGCGGCTGCGCTTCTGCACTCCCATGCTCTCGCA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 531 CTTTTGTGAGAGTCAAGCTTATTAAGCTGAGCTGAGACACTTCCCTTTTGA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 525 CTCTTATGCTGTCACAGATGTAATGAAGTTGCTTATGCAAGACCTTGGCCAAATGT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 591 CACCCCTCTCTGCTGTGCTCTTCACTTCTTCCCTTCCATCATCATGAGC
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 585 GGTATATGCTTACTGCAATCTGCTGTGCTGATGGGCTGAGCATTAATGTTCACTCTT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 651 CTCTATGCTTGAATCTAGGGGCTGTGCTCCGAATAGCTGTGCTGAGCGGAAAAA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 645 GTCCTATTTTCTGATTAATGAAGAGGCTTGCAGAGTCCCTTCAAGTCAAGAGGCGCAA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 711 AGCCCTGACCACTGCTCTCCCACTAAGAGCTGACCCCTCTTATGAGGGAGCCAT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 705 GGCCTTGGAACTGCTGTGACACATGAGTGTGTAATGCTGCTTCTATGTCACCTTAT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 771 GTTCATGT 778
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 765 TGGCTCT 772
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
  
```

RESULT 15
 US-08-465-980-1

; Sequence 1, Application US/08465980
 ; Patent No. 5756309
 ; GENERAL INFORMATION:
 ; APPLICANT: Soppet, Daniel R.
 ; APPLICANT: Li, Yi
 ; APPLICANT: Rosen, Craig A.

```

; APPLICANT: Ruben, Steven M.  

; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70  

; NUMBER OF SEQUENCES: 8  

; CORRESPONDENCE ADDRESS:  

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  

; STREET: 6 Becker Farm Road  

; CITY: Roseland  

; STATE: New Jersey  

; COUNTRY: USA  

; ZIP: 07068  

; COMPUTER READABLE FORM:  

; MEDIUM TYPE: Floppy disk  

; COMPUTER: IBM PC compatible  

; OPERATING SYSTEM: PC-DOS/MS-DOS  

; SOFTWARE: Patent Release #1.0, Version #1.30  

; CURRENT APPLICATION DATA:  

; APPLICATION NUMBER: US/08/465,980  

; FILING DATE: 06-JUN-1995  

; CLASSIFICATION: 536  

; ATTORNEY/AGENT INFORMATION:  

; NAME: Ferraro, Gregory D.  

; REGISTRATION NUMBER: 36,134  

; REFERENCE/DOCKET NUMBER: 325800-446  

; TELECOMMUNICATION INFORMATION:  

; TELEPHONE: 201-994-1700  

; TELEFAX: 201-994-1744  

; INFORMATION FOR SEQ. ID NO: 1:  

; SEQUENCE CHARACTERISTICS:  

; LENGTH: 1474 base pairs  

; TYPE: nucleic acid  

; STRANDEDNESS: single  

; TOPOLOGY: linear  

; MOLECULE TYPE: cDNA  

; FEATURE:  

; NAME/KEY: CDS  

; LOCATION: 274..1233  

; US-08-465-980-1
  
```

Query Match 12.9%; Score 122.4; DB 1; Length 1474;
 Best Local Similarity 49.0%; Pred. No. 8.1e-26;

Matches 327; Conservative 0; Mismatches 341; Indels 0; Gaps 0;

```

QY 111 AGTGGCCCTGTGGGAATGTCCTCCATCTTCCATCTACTGAGCGCTGACTTCA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 378 AGTGGCAATGTTGGAACTGATCGTGTCTTCATGTAAGAGGAGGACGACCTGCA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 171 CACCCCATGTAATCTTCTCCACGACCTCCCTCATGAGCTCATGTTGTTGTTAA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 438 CGCTCCATGTAATCTTCTCTGATGCTTGCAGCATGACCTGCTTATCCACATC
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 231 CATTGTCGCAAGAGGACCACTCTCTGTCGAGAGAGTCCATCTCTTGTGGG
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 498 CACCATGCTTAAGATCTGCTTCTGTTGATTCCGAGAGATTAAGCATTAAGG
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 291 CTGCGCATCAATATGGCTTTTGTCTCTCTGTGGATCTGAGGCGCTTGTGGG
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 558 CTGCTTACCAAGATGCTTCTTATTCATGCTTCAAGCATTAATCCACATCTGCT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 351 ACTGATGCTTATGACCACTAGGCGCTTACGACCACTTCACTATCCATCTCAT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 618 GGCATGCTCTTACCGCTTATGTCGATCTGCACTGCGCATGCTGAGTGT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 411 GAATCAGAGGCTGTCTCCAGATTACTGGAGCTCTGAGCGCTTGGGATTAATGATG
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 678 CAACATATACAGTACAGCCAGATTGCGATGCTGTGCTCGCGGATGCCCTTTT
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 471 AGTATTCAGATGCTGACGACCATGAGCTTACTTACTGTGCTCAAGAGCGTGATCA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 738 TTTTCCACCTCTGCTGATTAAGAGGCTGAGCTTCTGCACTCCATATCTCTGCGCA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 531 CTTTTCTGTGAGGTACAGCTTATTAAGAGCTGAGCTGTGACAGACTTCCCTTTTGA
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
  
```

```
Db 798 CTCTATTGTGTCCACGAGATGTAATGAAGTTGGCCATATGACACACTTGGCCCAATGT 857
QY 591 CACCCCTCTTTGCTTGTGCTGTGCTTCATGCTCTCCCTTCCCTTCATCATCATGAGC 650
    |||||
Db 858 GGTATATGTCTTACTGCCATTCTGTGTCATGCGGTGGACGTAATGTTCTCTT 917
    |||||
QY 651 CTCTATGCTTGCATCTAGGGGCTGTGCTCCGAATACGCTCTGCTCAAGGCTGGAAAAA 710
    |||||
Db 918 GTCTATTTTTCTGATAATACGAACGGTTCTGCACTGCCCTTCCAGTCAGAGCGGGCCAA 977
    |||||
QY 711 AGCCCTGGCCACCTGCTCTCCCTCCACTAACAGCTGTCAACCCCTCTATGCGGACGCCAT 770
    |||||
Db 978 GGCCTTGGAACTGTGTGTACACACATTGGTGTACTCGCCTTCTAATGTGCACTTAT 1037
    |||||
QY 771 GTTCATGT 778
    |||||
Db 1038 TGGCCTCT 1045
    |||||
```

Search completed: May 9, 2003, 04:49:05
Job time : 80 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2003, 04:31:20 ; Search time 1765 Seconds

(without alignments)
667.470 Million cell updates/sec

Title: US-09-634-109D-1

Perfect score: 948
Sequence: 1 atgggaagatgggtgaacca.....ggatgycagccagcactga 948

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | Length | DB | ID | Description |
|------------|-------|-------|--------|----|-------------------|-------------------|
| 1 | 935.4 | 98.7 | 1172 | 9 | US-10-023-601-7 | Sequence 7, Appl1 |
| 2 | 810.4 | 85.5 | 948 | 10 | US-09-886-055-176 | Sequence 176, App |
| 3 | 526.6 | 55.5 | 669 | 9 | US-10-259-430-19 | Sequence 19, Appl |
| 4 | 526.6 | 55.5 | 669 | 9 | US-10-259-430-27 | Sequence 27, Appl |
| 5 | 525 | 55.4 | 669 | 9 | US-10-259-430-23 | Sequence 23, Appl |
| 6 | 390.2 | 41.2 | 488 | 10 | US-09-747-155-212 | Sequence 212, App |
| 7 | 388.8 | 41.0 | 1290 | 10 | US-09-826-508-23 | Sequence 23, Appl |
| 8 | 361.8 | 38.2 | 975 | 10 | US-09-886-055-430 | Sequence 430, App |
| 9 | 360.6 | 38.0 | 951 | 10 | US-09-886-055-478 | Sequence 478, App |
| 10 | 345 | 36.4 | 948 | 10 | US-09-886-055-436 | Sequence 436, App |
| 11 | 343.8 | 36.3 | 1047 | 10 | US-09-886-055-434 | Sequence 434, App |
| 12 | 342 | 36.1 | 1110 | 10 | US-09-886-055-458 | Sequence 458, App |
| 13 | 339.8 | 35.8 | 954 | 10 | US-09-886-055-480 | Sequence 480, App |
| 14 | 330.2 | 34.8 | 1017 | 10 | US-09-886-055-428 | Sequence 428, App |
| 15 | 326.4 | 34.4 | 942 | 10 | US-09-886-055-164 | Sequence 164, App |
| 16 | 325.8 | 34.4 | 945 | 10 | US-09-886-055-144 | Sequence 144, App |
| 17 | 325.8 | 34.4 | 996 | 9 | US-09-898-586-17 | Sequence 17, Appl |
| 18 | 325.8 | 34.4 | 996 | 10 | US-09-761-288-17 | Sequence 17, Appl |
| 19 | 324.8 | 34.3 | 1047 | 10 | US-09-886-055-166 | Sequence 166, App |

| | | | | | | |
|----|-------|------|------|----|---------------------|-------------------|
| 20 | 324.2 | 34.2 | 994 | 9 | US-09-898-586-15 | Sequence 15, Appl |
| 21 | 324.2 | 34.2 | 994 | 10 | US-09-761-288-15 | Sequence 15, Appl |
| 22 | 320.4 | 33.8 | 996 | 10 | US-09-886-055-476 | Sequence 476, App |
| 23 | 313.4 | 33.1 | 939 | 10 | US-09-886-055-474 | Sequence 474, App |
| 24 | 308.4 | 32.5 | 1008 | 9 | US-09-974-591-13 | Sequence 13, Appl |
| 25 | 307.8 | 32.5 | 939 | 10 | US-09-886-055-422 | Sequence 422, App |
| 26 | 306.8 | 32.4 | 951 | 10 | US-09-886-055-198 | Sequence 198, App |
| 27 | 306.8 | 32.4 | 1008 | 9 | US-09-974-591-11 | Sequence 11, Appl |
| 28 | 303.2 | 32.0 | 1954 | 10 | US-09-864-761-5098 | Sequence 5098, Ap |
| 29 | 294.6 | 31.1 | 824 | 9 | US-09-989-442-75 | Sequence 75, Appl |
| 30 | 294.6 | 31.1 | 824 | 9 | US-10-073-865-64 | Sequence 64, Appl |
| 31 | 294.6 | 31.1 | 824 | 9 | US-10-103-313-230 | Sequence 230, App |
| 32 | 294.6 | 31.1 | 824 | 10 | US-09-764-853-351 | Sequence 351, App |
| 33 | 291.6 | 30.8 | 2282 | 10 | US-10-098-841-92 | Sequence 92, Appl |
| 34 | 291.6 | 30.8 | 2282 | 10 | US-09-747-835A-62 | Sequence 62, Appl |
| 35 | 290 | 30.6 | 1788 | 10 | US-09-747-835A-35 | Sequence 35, Appl |
| 36 | 290 | 30.6 | 2735 | 10 | US-09-747-835A-34 | Sequence 34, Appl |
| 37 | 285.2 | 30.1 | 916 | 9 | US-09-974-591-15 | Sequence 15, Appl |
| 38 | 284.2 | 30.0 | 1957 | 10 | US-09-864-761-13963 | Sequence 13963, A |
| 39 | 282.8 | 29.8 | 933 | 10 | US-09-864-761-30527 | Sequence 30527, A |
| 40 | 276.4 | 29.2 | 939 | 10 | US-09-886-055-66 | Sequence 66, Appl |
| 41 | 276.4 | 29.2 | 1050 | 9 | US-09-974-591-3 | Sequence 3, Appl1 |
| 42 | 276.4 | 29.2 | 1050 | 9 | US-09-974-591-7 | Sequence 7, Appl1 |
| 43 | 276.4 | 29.2 | 1050 | 9 | US-09-974-591-7 | Sequence 7, Appl1 |
| 44 | 275.4 | 29.1 | 939 | 10 | US-09-886-055-440 | Sequence 440, App |
| 45 | 274.8 | 29.0 | 1050 | 9 | US-09-974-591-9 | Sequence 9, Appl1 |

ALIGNMENTS

RESULT 1
US-10-023-601-7
Sequence 7, Application US/10023601
Publication NO. US20030083463A1
GENERAL INFORMATION:
APPLICANT: Padigarau, Muralidhara
APPLICANT: Kekuda, Ramesh
APPLICANT: Colman, Steven
APPLICANT: Spylek, Kimberly
APPLICANT: Ballinger, Robert
APPLICANT: Vernet, Corine
APPLICANT: Li, Li
APPLICANT: Shenoy, Suresh
APPLICANT: Casman, Stacie
APPLICANT: Guzey, Vladimir
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-224F
CURRENT APPLICATION NUMBER: US/10/023,601
CURRENT FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: 60/256,635
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/259,743
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: 60/299,327
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: 60/261,498
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/263,689
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/267,464
PRIOR FILING DATE: 2001-02-08
PRIOR APPLICATION NUMBER: 60/271,021
PRIOR FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: 60/275,946
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/278,150
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: 60/285,718
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 60/312,902
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 60/257,876

; PRIOR FILING DATE: 2000-12-21
 ; PRIOR APPLICATION NUMBER: 60/260,718
 ; PRIOR FILING DATE: 2001-01-10
 ; PRIOR APPLICATION NUMBER: 60/284,591
 ; PRIOR FILING DATE: 2001-04-18
 ; NUMBER OF SEQ ID NOS: 102
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO: 7
 ; LENGTH: 1172
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-023-601-7

Query Match 98.7%; Score 935.4; DB 9; Length 1172;
 Best Local Similarity 99.8%; Pred. No. 1.9e-287;
 Matches 947; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATGGGAAGATGGGTGACCAAGCTTACACAGATGGCTTCTCTTGGGCACTTTTCC 60
 DB 3 ATGGGAAGATGGGTGACCAAGCTTACACAGATGGCTTCTCTTGGGCACTTTTCC 62
 QY 61 CACAGCCAGACTGACCTTCTCTCTCTGAGTTATGGTGGCTTCCAGTGGCCCTC 120
 DB 63 CACAGCCAGACTGACCTTCTCTCTCTGAGTTATGGTGGCTTCCAGTGGCCCTC 122
 QY 121 TGTGGGAATGCTCCTCTCATCTCTCATCTACCTGAGCGTGCACACCCCATG 180
 DB 123 TGTGGGAATGCTCCTCTCATCTCTCATCTACCTGAGCGTGCACACCCCATG 182
 QY 181 TACTTCTCCAGAGCAGCTCCCTCATGAGACCTATGTTGGTGTGTACATTTGCCA 240
 DB 183 TACTTCTCCAGAGCAGCTCCCTCATGAGACCTATGTTGGTGTGTACATTTGCCA 242
 QY 241 AAGATGGCAAGCCAACTCTCTGTGCAAGAGTCCATCTCTTGTGGGCTGTGGCATA 300
 DB 243 AAGATGGCAAGCCAACTCTCTGTGCAAGAGTCCATCTCTTGTGGGCTGTGGCATA 302
 QY 301 CAAATTTGGCTTTTGTCTCTCTTGTGGATGTGAGGGCTTTGCTGGGACTCATGGCT 360
 DB 303 CAAATTTGGCTTTTGTCTCTCTTGTGGATGTGAGGGCTTTGCTGGGACTCATGGCT 362
 QY 361 TATGACCACTAGTGGCGGTTAGCCACCACCTTCACTATCCCATCATGTAATGAGG 420
 DB 363 TATGACCACTAGTGGCGGTTAGCCACCACCTTCACTATCCCATCATGTAATGAGG 422
 QY 421 GTCTGTCTCCAGATTAAGTGGAGCTCTGGGCTTTGGGATTAATGAGAGTGAATTAG 480
 DB 423 GTCTGTCTCCAGATTAAGTGGAGCTCTGGGCTTTGGGATTAATGAGAGTGAATTAG 482
 QY 481 ATGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGGATCATTTTCTGT 540
 DB 483 ATGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGGATCATTTTCTGT 542
 QY 541 GAGGTACAAAGCTTTAATGAAGTGGCGCTGTGAGACACTTCCCTTTTACACCCCTC 600
 DB 543 GAGGTACAAAGCTTTAATGAAGTGGCGCTGTGAGACACTTCCCTTTTACACCCCTC 602
 QY 601 TTGCTGTCTGTCTTATGCTTCTCTCTCTCTCATCATGATGAGCCCTCATGCT 660
 DB 603 TTGCTGTCTGTCTTATGCTTCTCTCTCTCTCATCATGATGAGCCCTCATGCT 662
 QY 661 TGCATC-CTAGGGGCTGTGCTCGAATAGCTCTGCTCAAGGCGTGGAAAAAGCCCTGGC 719
 DB 663 TGCATCCTAGGGGCTGTGCTCGAATAGCTCTGCTCAAGGCGTGGAAAAAGCCCTGGC 722
 QY 720 CACCTGTCTCTCCCACTAAGAGTGTCAACCTTCTATATGGGGAGGCACTGTTATGTA 779
 DB 723 CACCTGTCTCTCCCACTAAGAGTGTCAACCTTCTATATGGGGAGGCACTGTTATGTA 782
 QY 780 CCGTGAAGGCTTAGGCGCTACCGGGCCCTAGCCATGACAGAGTGGCGCTTATCTTACAC 839
 DB 783 CCGTGAAGGCTTAGGCGCTACCGGGCCCTAGCCATGACAGAGTGGCGCTTATCTTACAC 842

QY 840 ACTCCTTACTCCCATGCTGAACCCCTCATTTACAGTTGAGGAATGGGAGTATGGG 899
 DB 843 ACTCCTTACTCCCATGCTGAACCCCTCATTTACAGTTGAGGAATGGGAGTATGGG 902
 QY 900 GGCAGTGAAGAGGGGCTGGACCGCTGCAGAGTTGGCAGCCAGCACTGA 948
 DB 903 GGCAGTGAAGAGGGGCTGGACCGCTGCAGAGTTGGCAGCCAGCACTGA 951

RESULT 2 US-09-886-055-176

; Sequence 176, Application US/09886055
 ; Patent No. US2002013273A1

GENERAL INFORMATION:

; APPLICANT: STRYER, LUBERT
 ; APPLICANT: ZOULUVA, SERGEY
 ; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
 ; FILE REFERENCE: 078003-0277150
 ; CURRENT APPLICATION NUMBER: US/09/886,055
 ; PRIOR FILING DATE: 2001-06-22
 ; PRIOR APPLICATION NUMBER: 60/213,812
 ; NUMBER OF SEQ ID NOS: 522
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO: 176
 ; LENGTH: 948
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-886-055-176

Query Match 85.5%; Score 810.4; DB 10; Length 948;
 Best Local Similarity 90.9%; Pred. No. 1.1e-247;
 Matches 862; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1 ATGGGAAGATGGGTGACCAAGCTTACACAGATGGCTTCTCTTGGGCACTTTTCC 60
 DB 1 ATGGGAAGATGGGTGACCAAGCTTACACAGATGGCTTCTCTTGGGCACTTTTCC 60
 QY 61 CACAGCCAGACTGACCTTCTCTCTCTGAGTTATGGTGGCTTCCAGTGGCCCTC 120
 DB 63 CACAGCCAGACTGACCTTCTCTCTCTGAGTTATGGTGGCTTCCAGTGGCCCTC 120
 QY 121 TGTGGGAATGCTCCTCTCATCTCTCATCTACCTGAGCGTGCACACCCCATG 180
 DB 123 TGTGGGAATGCTCCTCTCATCTCTCATCTACCTGAGCGTGCACACCCCATG 180
 QY 181 TACTTCTCCAGAGCAGCTCCCTCATGAGACCTATGTTGGTGTGTACATTTGCCA 240
 DB 183 TACTTCTCCAGAGCAGCTCCCTCATGAGACCTATGTTGGTGTGTACATTTGCCA 240
 QY 241 AAGATGGCAAGCCAACTCTCTGTGCAAGAGTCCATCTCTTGTGGGCTGTGGCATA 300
 DB 243 AAGATGGCAAGCCAACTCTCTGTGCAAGAGTCCATCTCTTGTGGGCTGTGGCATA 300
 QY 301 CAAATTTGGCTTTTGTCTCTCTTGTGGATGTGAGGGCTTTGCTGGGACTCATGGCT 360
 DB 303 CAAATTTGGCTTTTGTCTCTCTTGTGGATGTGAGGGCTTTGCTGGGACTCATGGCT 360
 QY 361 TATGACCACTAGTGGCGGTTAGCCACCACCTTCACTATCCCATCATGTAATGAGG 420
 DB 363 TATGACCACTAGTGGCGGTTAGCCACCACCTTCACTATCCCATCATGTAATGAGG 420
 QY 421 GTCTGTCTCCAGATTAAGTGGAGCTCTGGGCTTTGGGATTAATGAGAGTGAATTAG 480
 DB 423 GTCTGTCTCCAGATTAAGTGGAGCTCTGGGCTTTGGGATTAATGAGAGTGAATTAG 480
 QY 481 ATGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGGATCATTTTCTGT 540
 DB 483 ATGTGGCAGCCATGGGCTTACCTTACTGTGGCTCAAGGAGCGTGGATCATTTTCTGT 540
 QY 541 GAGGTACAAAGCTTTAATGAAGTGGCGCTGTGAGACACTTCCCTTTTACACCCCTC 600
 DB 543 GAGGTACAAAGCTTTAATGAAGTGGCGCTGTGAGACACTTCCCTTTTACACCCCTC 600

Db 541 GAGATGCTATCTTGTGAGCTGGCCCTGTGTGACACATCCCTTTTGAGAGGTGATA 600
Qy 601 TTGCTGCTGTGTCTTCATGCTTCCCTTCCATCATGATGAGCTGATGCT 660
Db 601 TTGCTGCTGTGTCTTCATGCTTCCCTTCCATCATGATGATGAGCTGATGCT 660
Qy 661 TGCATCTGAGGGGCTGTGTCTCGAATAGCGTCTGCTCAGGCTTGAAAAAGCCCTGGCC 720
Db 661 CACATTTAGAGGAGCTGTGTGCAAAATGCACTGTGCTCAGGCTTGAAAAAGCCCTGGCC 720
Qy 721 ACCCTGCTCTCCACCTTAACAGCTGTCACTCTTCTATGAGGAGCCATGTTCATGAC 780
Db 721 ACCCTGCTCTCCACCTTAACAGCTGTCACTCTTCTATGAGGAGCCATGTTCATGAC 780
Qy 781 CTGAGGCGCTAGGCGCTAGCGGCGCCCTAGCCATGACAGAGTGGCCCTATCTTACACA 840
Db 781 CTGAGGCGCTAGGCGCTAGCGGCGCCCTAGCCATGACAGAGTGGCCCTATCTTACACA 840
Qy 841 GTCTTTACTCCCATCTGTAACCCCTCATTTTACAGCTTGAGGAATGGAGGTGATGGGG 900
Db 841 GTCTTTACTCCCATCTGTAACCCCTCATTTTACAGCTTGAGGAATGGAGGTGATGGGG 900
Qy 901 GCACGTAGGAGGGGCTGAGCCGCTGACAGAGATGGGAGCCAGCACTGA 948
Db 901 GCACGTAGGAGGGGCTGAGCCGCTGACAGAGATGGGAGCCAGCACTGA 948

RESULT 3
US-10-259-430-19

; Sequence 19, Application US/10259430
; Publication No. US20030082615A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Randall
; APPLICANT: Yau, King-Wai
; APPLICANT: Krautwurst, Dietmar
; TITLE OF INVENTION: Olfactory Receptor Expression Libraries
; TITLE OF INVENTION: ad Methods of Making and Using Them
; FILE REFERENCE: 001107.00105
; CURRENT APPLICATION NUMBER: US/10/259.430
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/465.901
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/112,605
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: PCR primer
US-10-259-430-19

Query Match 55.5%; Score 526.6; DB 9; Length 669;
Best Local Similarity 86.7%; Pred. No. 2,1e-157;
Matches 580; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 193 AGCCAGCTCTCCCTCATGACCTCATGTTGCTGTATACATTTGSCCAAAAGATGACACC 252
Db 1 AGTCAGCTCTCCCTCATGACCTCATGTTGCTGTATACATTTGSCCAAAAGATGACACC 60
Qy 253 AACTTCCTGTCTGCGAGAGATCATCTCTTGTGGGCTGTGGCATCAAAATGGCTTT 312
Db 61 AACTTCCTGTCTGCGAGAGATCATCTCTTGTGGGCTGTGGCATCAAAATGGCTTT 120
Qy 313 TTTGCTCTCTTGTGGATCTGAGGGCTCTTGTGGGACTCATGGCTTATGACACTAC 372
Db 121 TTTGCTCTCTTGTGGATCTGAGGGCTCTTGTGGGACTCATGGCTTATGACACTAT 180
Qy 373 GTGGCGTTAGACACCCATTCATATCCATCCATCCATGAATGAGGGTGTCTCAG 432
Db 181 GTGGCGTTAGACACCCATTCATATCCATTCATATGAGCCAAAGGTCTGTCTCAG 240

Qy 433 ATTATGGAGGCTCCTGGGCTTTGGGATATAGATGAGATTCAGATGCTGGACGC 492
Db 241 ATTGCTGGAAGTTCCTGGGCTTTGGGATTCATGATGATATTAATACATGCTGGACGC 300
Qy 493 ATGGGCTTACTTACTGTGTGCTCAAGAGCGTGAATCATCTTTTCTGTGAGGTACAGCT 552
Db 301 ATGAGCTGCGCTACTGTGTGCTCAAGAGCGTGAATCATCTTTTCTGTGAGGTACAGCT 360
Qy 553 TTATGGAAGCTGCGCTGTGAGACACTTCCCTTTTGACACCCCTCTTCTGCTGCT 612
Db 361 TTATGGAAGCTGCGCTGTGAGACACTTCCCTTTTGACACCCCTCTTCTGCTGCT 420
Qy 613 GTCTTCATCTCTCTCCCTTCCCTTTCATCATGAGGCTCTTATGCTTGCATCCAGG 672
Db 421 GTCTTCATCTCTCTCCCTTCCCTTTCATCATGAGGCTCTTATGCTTGCATCCAGG 480
Qy 673 GCTGTCTCCGAATAGCTCTGTGCTCAGGCGCTGGAAGGAGCCGCTGACCTGTCTCC 732
Db 481 GCTGTCTCCGAATAGCTCTGTGCTCAGGCGCTGGAAGGAGCCGCTGACCTGTCTCC 540
Qy 733 CACCTAAGAGCTGACCCCTTCTATGAGGAGCCATGTCATGCTTACAGTCTTACTCC 792
Db 541 CACCTAAGAGCTGACCCCTTCTATGAGGAGCCATGTCATGCTTACAGTCTTACTCC 600
Qy 793 CGCTACCGGGCCCTAGGCAATGACAGAGTGGGCTCTATCTTACAGTCTTACTCC 852
Db 601 CGATATCGGCGCTCCAGCATGACAAAGTGTCTCATCTTACAGTCTTACTCT 660
Qy 853 ATGCTGAC 861
Db 661 ATGCTGAC 669

RESULT 4

US-10-259-430-27
; Sequence 27, Application US/10259430
; Publication No. US20030082615A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Randall
; APPLICANT: Yau, King-Wai
; APPLICANT: Krautwurst, Dietmar
; TITLE OF INVENTION: Olfactory Receptor Expression Libraries
; TITLE OF INVENTION: ad Methods of Making and Using Them
; FILE REFERENCE: 001107.00105
; CURRENT APPLICATION NUMBER: US/10/259.430
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/465.901
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/112,605
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: PCR primer
US-10-259-430-27

Query Match 55.5%; Score 526.6; DB 9; Length 669;
Best Local Similarity 86.7%; Pred. No. 2,1e-157;
Matches 580; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 193 AGCCAGCTCTCCCTCATGACCTCATGTTGCTGTATACATTTGSCCAAAAGATGACACC 252
Db 1 AGTCAGCTCTCCCTCATGACCTCATGTTGCTGTATACATTTGSCCAAAAGATGACACC 60
Qy 253 AACTTCCTGTCTGCGAGAGATCATCTCTTGTGGGCTGTGGCATCAAAATGGCTTT 312
Db 61 AACTTCCTGTCTGCGAGAGATCATCTCTTGTGGGCTGTGGCATCAAAATGGCTTT 120

Query Match 41.2%; Score 390.2; DB 10; Length 488;
Best Local Similarity 88.9%; Pred. No. 4,8e-114;
Matches 433; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

QY 373 GTGGCGGTAGCCACCCCTTACATATCCCATCTCATGAAATCAGAGGCTGTCTCCAG 432
DB 2 GTGGCGGTAGCCACCCCTTACATATCCCATCTCATGAAATCAGAGGCTGTCTCCAG 61
QY 433 ATTAATGGAGAGCTCTGGGCGCTTTGGATTAATAGATGG-AGTGAATTCAGATGGTGGCAGC 491
DB 62 ATTAACGGGAGAGCTCTGGGCGCTTTGGATTAATAGATGGTGAATTCAGATGGTGGTGTGT 121
QY 492 CATGGGCTTACCTTACTGTGGCTCAAGGAGCGGTGATCACTTTTCTGTGAGTACAAGC 551
DB 122 AATGAATTTCCCTTACTGTGGCTTGAAGAGGTGAACCATTTCTTCTGTGAGATCTAATC 181
QY 552 TTTAATGAAGCTGGGCTGTGACAGACTTCCCTTTTACACACCTTCTTGTCTGTCTG 611
DB 182 CTGTGTGAAGCTGGGCTGTGAGACACATCCCTGTGTGAGAGGTGATATTTGTCTGTG 241
QY 612 TGTCTTCAATGCTTCTTCTTCCCTTCTCATATCATATGAGCGCTTCTGTGATCTTACTAG 671
DB 242 TGTCTTCAATGCTTCTTCTTCCCTTCTCATATCATATGAGCGCTTCTGTGATCTTACTAG 301
QY 672 GCGTGTGCTCGGAATACGCTGTGCTCAGGCGCTGAGAAAAAGCCCTGGCCACCTGCTCTC 731
DB 302 GACTGTGCTGCAAAATGCACTGTGCTCAGGCGCTGAGAAAAAGCCCTGGCCACCTGCTCTC 361
QY 732 CCACCTAAGAGCTGTACACCTCTTCTATGAGGCGAGCCATGTTTCAATGATACGTGAGGCTAG 791
DB 362 CCACCTAAGAGCTGTACACCTCTTCTATGAGGCGAGCCATGTTTCAATGATACGTGAGGCTAG 421
QY 792 GCGGTACCGGGGCGGCTTACAGAGGTGGGCTTCTTCTTCTTCAATGATACGTGAGGCTAG 851
DB 422 GCGGTACCGGGGCGGCTTACAGAGGTGGGCTTCTTCTTCTTCAATGATACGTGAGGCTAG 481
QY 852 CATGCTG 858
DB 482 CATGCTG 488

RESULT 7
US-09-826-508-23
; Sequence 23, Application US/09826508
; Patent No.: US20010025099A1
; GENERAL INFORMATION:
; APPLICANT: Nabil Elshourbagy
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
; FILE REFERENCE: GP-7074USB
; CURRENT APPLICATION NUMBER: US/09/826,508
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 1290
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-826-508-23

Query Match 41.0%; Score 388.8; DB 10; Length 1290;
Best Local Similarity 64.4%; Pred. No. 2.1e-113;
Matches 582; Conservative 0; Mismatches 322; Indels 0; Gaps 0;

QY 9 ATGGGTGACACAGCTTCAACAGATGCTTCTTCTTGGGCACTTTTCCACAGCCCA 68
DB 301 AAGGAGATTCGACCTTCACTCGACTTCACTTCTTGGGAACTTCAATCAACAGCC 360
QY 69 GAGTACCTTGTCTTCTTCTGAGTATGAGTGTTCACAGGCGGCTGTGGGAA 128
DB 361 CACCCACACCTTCTTCTTCTTCTGAGTATGAGTGTTCACAGGCGGCTGTGGGAA 420

QY 129 TGTCTTCAATGCTTCTTCTTCCCTTCTCATATCATATGAGCGCTTCTGTGATCTTACTTCT 188
DB 421 CTCTGTCAATGCTTCTTCTTCCCTTCTCATATCATATGAGCGCTTCTGTGATCTTACTTCT 480
QY 189 CCTACGACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 248
DB 481 CCTACGACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
QY 249 AGCCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 308
DB 541 CTTCACATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600
QY 309 CTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 368
DB 601 CTCTATATCATATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 660
QY 369 CTACGTGGCGCTTACGACCCCTTCACTATCCATCTCTCATGAATCAAGAGGCTGTCT 428
DB 661 CTACGTGGCGCTTACGACCCCTTCACTATCCATCTCTCATGAATCAAGAGGCTGTCT 720
QY 429 CCAGATTAAGTGGGAGCTGCTGGGCGCTTGGGATTAATGAGTGAATTCAGATGGTGGC 488
DB 721 ACTTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 780
QY 489 AGCCATGAGGCTTACCTTACTGTGCTCAAGAGCGGTGATCACTTTTCTGTGAGTACA 548
DB 781 CACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840
QY 549 AGCTTATTAAGAGTGGGCTGTGACAGACACTTCCCTTTTTCACACCCCTCTTCTTCT 608
DB 841 TTCCCTACTAATCTCTCTCATGCAATGAGACACATCAATATTTGAAAGTATTTTCAATTTG 900
QY 609 CTGTCTTCAATGCTTCTTCTTCCCTTCTCATATCATATGAGCGCTTCTGTGATCTTACT 668
DB 901 CTCTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 960
QY 669 AGGGGCTGTCTCGAATACGCTGTGCTCAGGCGCTGAGAAAAAGCCCTGGCCACCTGTCT 728
DB 961 TGTGCTGTCAATTCACATGAGGATCTGAGAGGCGTGTGAGAAAAAGCTTTCACAGCCTGTT 1020
QY 729 CTCCACCTAAGAGCTGTACACCTCTTCTATGAGGCGAGCATGTTTCAATGATACGTAG 788
DB 1021 CTCTACCTCATGAGTGTGAGGATGTTCTATGAGGAGGAGTGTTCATGATATCATACAGC 1080
QY 789 TAGGGCTTACCGGCGGCTTACAGAGGTGGGCTTCTATCTTCAACAGTCTTAC 848
DB 1081 CACATCTATGCTCTCCCAACAGCAGAGAGAGTGTGTGTATTTCTACACCATCTTCTAC 1140
QY 849 TCCCATGTGAACCCCTCATTTTACAGTGTGAGAAATGGGAGGTGATGGGCGACGTAG 908
DB 1141 TCCCATGTGAATCCCTCATCTACAGGCTCGCAACAGAAATGACAGACATTCAT 1200
QY 909 GAAG 912
DB 1201 GAAG 1204

RESULT 8
US-09-886-055-430
; Sequence 430, Application US/09886055
; Patent No.: US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 430
LENGTH: 975
TYPE: DNA
ORGANISM: Homo sapiens
US-09-886-055-430

Query Match 38.2%; Score 361.8; DB 10; Length 975;
Best Local Similarity 62.6%; Pred. No. 7.2e-105;
Matches 564; Conservative 0; Mismatches 337; Indels 0; Gaps 0;

26 ACACGAGTGGCTTCTCTCTGGGCACTTTCCACAGCCAGCACTACCTTGCTCT 85
26 ACTCACAACCTTCTCTCTGGGCACTTTCCACAGCCAGCACTACCTTGCTCT 85
86 TCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 145
86 TTGCAATAGTCT 145
146 TCATCTACCTGAGCGCTGAGCTTACAGCCCGCATCTCTCTCTCTCTCTCT 205
146 TCATCTACCTGAGCGCTGAGCTTACAGCCCGCATCTCTCTCTCTCTCTCT 205
206 TCATGAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 265
206 TCATGAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 265
266 GCAGAGAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 325
266 AGGACAACACCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 325
326 TGGATCTGAGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 385
326 TTGGGAGGGAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 385
386 ACCCACTCACTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 445
386 ACCCTCTACGATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 445
446 CCGGAGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 505
446 CCGGAGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 505
506 ACTGAGCTCAAGAGCGCTGATCCTTTTCTCTCTCTCTCTCTCTCTCTCT 565
506 TCTGAGATATCCGAGAGATCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 565
566 CCGGAGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 625
566 CCGGAGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 625
626 TTATCT 685
626 TTATCT 685
686 TACGCT 745
686 TACGCT 745
746 TCACCT 805
746 TCACCT 805
806 CAGAGAAAGATTAAGTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 865
806 CAGAGAAAGATTAAGTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 865
866 TCATCTACCTGAGCGCTGAGCTTACAGCCCGCATCTCTCTCTCTCTCT 925
866 TCATCTACCTGAGCGCTGAGCTTACAGCCCGCATCTCTCTCTCTCTCT 925
926 G 926
926 G 926

RESULT 9
US-09-886-055-478
Sequence 478, Application US/09886055
Patent No. US20020132273A1
GENERAL INFORMATION:
APPLICANT: STREYER, LOBERT
APPLICANT: ZOZULA, SERGEY
TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
FILE REFERENCE: 078003-0277150
CURRENT APPLICATION NUMBER: US/09/886,055
CURRENT FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: 60/213,812
PRIOR FILING DATE: 2000-06-22
NUMBER OF SEQ ID NOS: 522
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 478
LENGTH: 951
TYPE: DNA
ORGANISM: Homo sapiens
US-09-886-055-478

Query Match 38.0%; Score 360.6; DB 10; Length 951;
Best Local Similarity 62.7%; Pred. No. 1.7e-104;
Matches 561; Conservative 0; Mismatches 334; Indels 0; Gaps 0;

33 TGGCTTCT 92
21 TGACTTACCT 80
93 AGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 152
81 GATCCTGCT 140
133 CCGGAGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 212
141 GGTGAGCT 200
213 CCGGAGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 272
201 CACCTTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 260
273 GTCACT 332
261 GATCCTGCT 320
333 TGGGAGGCT 392
321 TGAGTTCCT 380
393 TCATCTACCT 452
381 GAGATACCT 440
453 CTTTGGAGTATAGATGAGTGAATGATGATGATGATGATGATGATGATGAT 512
441 TGGGAGGCT 500
513 CCGGAGTGGGCT 572
501 CTCCTGAGTATCAACCATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 560
573 AGACACTTCT 632
561 AGACAGTCT 620
633 CTTTCT 692
621 CATCT 680
693 TGCTTACGCTCTGAGAAAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 752


```

Db 681 TCTTAAGGTCGCAAAAAGGCGCTTACACACTGTTCCCTCCACCTGAGCTGATAGAT 740
Oy 753 CTTCATGAGGAGGACATGTTTCATGTTACCTGAGGCTGAGGCGTACCGGGCCAGGCA 812
Db 741 CTTCATGAGGAGGCTGCTTCTACACATACGTGCTGCCCAAGTCCTTCCACACCCCGACA 800
Oy 813 TGACAAGGTGGCTCTATCTCTACACAGTCTTACTCCCATGCTGAAGCCCTCATTTA 872
Db 801 GGACAAAGTAGTGAAGCTTCTATACCATTTGTCAGGCCCATGCTTAATCCCTCATCTA 860
Oy 873 CAGCTTGAGGAATGGGAGGTATGGGGCAGTGAAGAGGGCTGACCCGCTGC 927
Db 861 CAGCCTGAGAAACAAGAGGCTCATAGGGGCAATTTAAAGATTTGTCATGTGC 915

RESULT 10
US-09-886-055-436
; Sequence 436, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STREYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSOR PERCEPTION, AND
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886, 055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213, 812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 436
; LENGTH: 948
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-886-055-436

Query Match 36.4%; Score 345; DB 10; Length 948;
Best Local Similarity 61.3%; Pred. No. 1.6e-99;
Matches 555; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

Oy 7 AGATGGTGAACCAAGTCTTACACAGAGGCTTCTCTCTGGGCACTTTTCCACAGC 66
Db 16 AGGATGGCCACACACAGTGAAGTTGATTTCACTCATGGAGCTCTTACAGCATCC 75
Oy 67 CAGAGTGAAGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 126
Db 76 AAGCATCCAGCTCTACTAGTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGA 135
Oy 127 AATGCTCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 186
Db 136 AATGCTCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 195
Oy 187 TTCTCAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 246
Db 196 TTCAATCAATCAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 255
Oy 247 GCAGCAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 306
Db 256 CTCCTGGACAGGTGATGGGTGTAATAGGTCTGAGCCCTGATGGGATCAGATG 315
Oy 307 GCGTTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 366
Db 316 TTCTCTATCTGACATGACAGGTTGGAATTTTCTCTCTCTCTCTCTCTCTCTCTCT 375
Oy 367 CACTACGTCGCGCTTACGACCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 426
Db 376 GGTACGTCGCGCTTACGACCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 435
Oy 427 CTCAGATTAAGTGGAGGCTCTGGGCTTTGGGATTAATAGATGAGATGATTCATG 486
Db 436 CTTTCTCTGAGTGGGCTCTGCTGTCTCTGAGGTGATGAGGTCTCTCTCTCTCTCT 495

```

```

Oy 487 GCAGCATAGGCTTACCTTACTGTGCTCAAGAGCGTGATACATTTTCTGAGGTA 546
Db 496 ATACCATAGAGCTTCCCTTCTGAGATCCGAGGATTCATCATTTCTCTGTAAGTC 555
Oy 547 CAGCTTATTTGAAGTGGCTGTGTCAGACACTTCTCTTTTGTAGACCCCTCTCTCT 606
Db 556 CCGTGTGTAAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 615
Oy 607 TGCTGTCTCTCATAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 666
Db 616 TGCTGTCTCTCATAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 675
Oy 667 CTAGGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 726
Db 676 CTCCTACCGTCTCACAGAGATGATACACAGAGGCGGAAAAAGGCTTTGCCACCTGC 735
Oy 727 TCTTCCACCTAACAGCTGTACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 786
Db 736 TCTTCCACCTGTACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 795
Oy 787 CCTAGGCGCTTACCGGCGCTTACAGCATGACAGGATGAGGCTCTCTCTCTCTCTCT 846
Db 796 CCGAGCTCTCTACACACCCCTGAGAGGATGATGTCTCTCTCTCTCTCTCTCTCTCT 855
Oy 847 ACTCCCATGCTGAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 906
Db 856 ACTCCGCTCTGMAACCTTTATCTATGTCCTTAAGATTAAGATGATGAGGCTCTCT 915
Oy 907 AGGAA 911
Db 916 AAGAA 920

```

```

RESULT 11
US-09-886-055-434
; Sequence 434, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STREYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSOR PERCEPTION, AND
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886, 055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213, 812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 434
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-886-055-434

Query Match 36.3%; Score 343.8; DB 10; Length 1047;
Best Local Similarity 62.1%; Pred. No. 4e-99;
Matches 543; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

Oy 37 TTCTCTCTCTTGGGCACTTTTCCACAGCCAGACGCTGTCTCTCTCTCTCTCTCTCT 96
Db 130 TTCAATCAATCAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 189
Oy 97 ATGTGTCTCTTACAGTGGGCTCTGTGGGATGTCTCTCTCTCTCTCTCTCTCTCTCT 156
Db 190 TTGTGTGTCTCTGATGGGCTGTCTGTGAATGCTGTCTGTCTCTCTCTCTCTCTCT 249
Oy 157 GAGCTGAGCTTACAGCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 216
Db 250 GAGCCCATCTCTACACCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 309
Oy 217 ATGTGTCTGTAACTTTGCCAAAGATGACCACTTCTCTCTCTCTCTCTCTCTCTCTCT 276

```

Db 310 GCGTACATTTGTCACGTGCGCCAGAGATGCTCTGAGACCAAGTCATGAGGTGTAATAG 369
Qy 277 ATTCCTTTGGGCGTGGGATACAAATTTGGCTTTTGTCTCTTGTGGGATCAG 336
Db 370 ATCTAGCCCTGAGTGGGATGAGATGTTCTTCTACGTACCTAGCAGGTTCAAG 429
Qy 337 GGGCTTTGCTGGGATGAGCTTATGACCACTACGTCGCGCTTGGCCACTTAC 396
Db 430 TTTTCTTTAGCCACCATGAGCCCTACGCGCTACGTCGATCTGCGCTTCCGT 489
Qy 397 TATCCATCTTCATGAAACAGAGGCTGTCTCCAGATTACTGGAGCTCTGGGCTTT 456
Db 490 TACCTGCTTCATGAAACCATGAGGCTGTCTCTTCTGTCATGAGGCTGCTCTG 549
Qy 457 GGGATATAGATGAGATGATTCAGATGAGGCTGAGGCTGAGGCTTACTGAGCTCA 516
Db 550 GCGTACGATGAGGCTTACATTCACCTCCATCACCATGACCTTCCCTTCCGTGATCC 609
Qy 517 AGAGCGTGATCATTCTTTCTGTAGGTACAAAGCTTATGAAAGCTGAGCTGACAG 576
Db 610 CGGAGATTTCATTTCTTCTGTAGAGTCTCTGCTATTGAATCTCTCCGCTGACAG 669
Qy 577 ACTTCCCTTTTGACACCTCTCTTGTGCTGTCTGTCTGATCTCTCTCTCTCTCT 636
Db 670 ACCCTACTCTATGAGATTTTCTATGATCTGCTGTCTGTCTCTCTCTCTCTCTCTG 729
Qy 637 TCCATCATCATGAGCTCTCTATGCTTGCATCTGAGGCTGCTGCTGCTGCTGCTG 696
Db 730 GTGATCATTTCAAGCTCTCTATTTACTCATCTCTCTCTCTCTCTCTCTCTCTCTG 789
Qy 687 CAGGCTTGAAAAAGCCCTGGCCACCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 756
Db 790 GAGGCGCGGAAAAAGGCTTTGGCCACCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 849
Qy 757 TATGGGCGAGCATGTCATGTCATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 816
Db 850 TATGGGCGTGCATCTACACTACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 909
Qy 817 AAGGTGCGCTTATCTTCTACACAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 876
Db 910 ATGATGATCTGCT 969
Qy 877 TTGAGGATGGGAGGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 911
Db 970 CTAGGATTAAGGATGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1004

RESULT 12

US-09-886-055-458
; Sequence 458, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STREYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886, 055
; PRIOR APPLICATION NUMBER: 2001-06-22
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 458
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-886-055-458

Query Match 36.1%; Score 342; DB 10; Length 1110;
Best Local Similarity 60.8%; Pred. No. 1,5e-98;
Matches 558; Conservative 0; Mismatches 360; Indels 0; Gaps 0;

Qy 13 GTGAACAGTCTACACAGATGAGCTTCTCTCTGAGCATCTTTCCACAGCCAGACT 72
Db 160 GAGTACACACATCTCTACAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 219
Qy 73 GACCTTGCT 132
Db 220 TCAGCTCTTATTTTGGCCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 279
Qy 133 CTCTCATCT 192
Db 280 GTATGATCT 339
Qy 193 AGCCAGCT 252
Db 340 AGCCAGCT 399
Qy 253 AACTTCT 312
Db 400 AATTAACCTGATGATCAAGAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 459
Qy 313 TTTGCT 372
Db 460 TACCTTACCT 519
Qy 373 GTGGCGTTAGCCACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 432
Db 520 GTGGCGATTTGCAACCTCTGAGATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 579
Qy 433 ATTATGAGGACT 492
Db 580 ATTATGAGGACT 639
Qy 493 ATGGCTTACCTTCT 552
Db 640 ATGAGCTTCT 699
Qy 553 TTATGAGGCT 612
Db 700 GTCCGAAAGTTGGCATGTGAGACACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 759
Qy 613 GTCTTCACT 672
Db 760 GTTTTGAATGCT 819
Qy 673 GCTGTCTCCAAATACGCT 732
Db 820 ACAGTTACGTGATGAGCT 879
Qy 733 CACCTACAGCT 792
Db 880 CACATGACTGTGTCT 939
Qy 793 CGCTACGGGGGCGCTAGCATGACAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 852
Db 940 TCTTACCAACAGCCAGCCAGGACAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 999
Qy 853 ATGCTGAACCCCT 912
Db 1000 ATGCTGAACCCCT 1059
Qy 913 GGGCTGACCCCTGACG 930
Db 1060 GCGTTGGGAGGTTCAAG 1077

RESULT 13

US-09-886-055-480
; Sequence 480, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STREYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY

```

; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; FILE REFERENCE: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-027150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 480
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-886-055-480

```

```

Query Match      35.8%; Score 339.8; DB 10; Length 954;
Best Local Similarity 60.8%; Pred. No. 7.2e-98;
Matches 554; Conservative 0; Mismatches 357; Indels 0; Gaps 0;

```

```

QY 16 AACGAGCTACACAGATGGCTTCTTCTGCGATCTTTTCCACAGCCAGACTGAC 75
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13 AATATATCCGTATGCGGACCTTATCCCTTGCGGTGTTCAGCAACGCCGTTCCCC 72
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 76 CTTGTCTCTTCTGAGTTATGTTGTTCTTACAGTGGCCCTCTGTGGAAATGCTC 135
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73 TGGCTTCTCTTCCCTCATCTCCGCTCTTTTGAACCTCCATAGCCACAGCTGCTC 132
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 136 CTCATCTTCTCATCTACCTGAGCGTGGACTTCACACCCCCCATGCTCTTCAGC 195
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 133 AAGATCATTTTCATCCCATATAGACTCCCGCTCACACCCCCCATGCTCTTCAGC 192
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 196 CAGCTCTCCCTCATGACCTCATGTTGCTGTAACTTGTGCCAAGATGGACGCCAC 255
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 193 CAGCTCTCCCTCATGACCTCATGTTGCTGTAACTTGTGCCAAGATGGACGCCAC 252
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 256 TTCTGTCTGGAGAGTCCATCTCTTGTGGGCTGGGCTGCATACAAATGGCTTTT 315
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 253 CAGGTGTGAGCCAGAGAGCAATTCCTTGTGATGCACTGCCCAACACTTCTCTAC 312
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 316 GCTCTCTTGTGGAGTCTGAGGGCTCTTGTGAGCTCATGCTTATGACCACTAGC 375
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 313 TTGACCTTACAGAGGGGTGAGTTCTTCTCTAGGACTCATGCTCTATGATCCGTAC 372
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 376 GCCGTATGACCCACTTCTCATATCCATCCATGATGAGAGGCTCTGCTCCAGATT 435
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 373 GCCATCTGCAACCTCTGCACTATCCGCTCATGAGCGGCAAGATCTCGTTGATT 432
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 436 ACTGGAGCTCTGGGCTTTGGATATATGATGAGTATTAGATGTTGAGAGCCATG 435
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 433 GTGGCGGACGCTGGGCTGGAGGCTATGATGATTCTTGTCTACCCCCGTCACCATG 432
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 496 GGCCTACCTTACTGTGCTCAAGAGGTGATCACTTTTGTGTGATGACCAAGCTT 555
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 493 CAGTTCCCTCTGCTGCTCTGCGGAGATCAACCACTTCTTGTGCGAGTGCTGCTT 552
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 556 TTGAAGCTGGCTGTGACAGACTTCCCTTTTGTGACACCCCTCTTGTGCTGTGTC 615
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 553 CTGAAGCTCTCTGCAAGGACACATAGCTTACGAGACCATGATGATGCTGTGATT 612
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 616 TTGATGTTCTCTTCTCTCTCTCATATGATGAGGCTCTCTATGCTTGTGATGGGCT 675
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 613 ATATGCTCTCATCCCTTCTCTGATCTGAGGCTCTTACCAAGAAATGCTATTAAT 672
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 676 GTGCTCCGATATAGCTCTGCTCAGGCTGGAAAAAGCCCTGCGACCTCTCTCCAC 735
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 673 GTTATATAGATGAGCGAGAGGAGGAGGAAAGCTTGGCCACCTCTCTCTACAC 732
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 736 CTAAACAGCTGTCACTCTTCTATGAGGAGCCATGTTATGATCTGAGGCTTAGCGC 795
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 733 ATGATGTTGTGAGCTCTCTCTATGAGGAGGCTGATGATACATATAGCTGCTCATTT 792
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 796 TACCGGGCCCTGAGCATGATGAGGCTGCTCTATCTTACACAGATCTTACTCCATG 855
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 793 TACCACACCCCTGAGCAGCAAGCTATCTGCTTACACCATCTTACTCCATG 852
QY 856 CTGAACCCCTCATTTTACAGCTTGAAGATGGAGGTGAGGAGGACTGAGAAAGG 915
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 853 CTGATCTACTCATTTTACAGCTTGAAGATGATGATGATGATGATGATGATGAT 912
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 916 CTGAGACCGCTG 926
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 913 GTGGGAGGTG 923
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 14
US-09-886-055-428
; Sequence 428, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY

```

```

; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; FILE REFERENCE: BIOSENSORS OF CHEMICAL SENSANTS
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 428
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-886-055-428

```

```

Query Match      34.8%; Score 330.2; DB 10; Length 1017;
Best Local Similarity 60.8%; Pred. No. 8.5e-95;
Matches 539; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

```

```

QY 40 TTCTCTTGGGCACTTTTCCACAGCAGCACTGACCTTCTCTCTGCAATTATG 99
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 100 TTATTCCTTGGGTTGTTGAGACGCCGTTTCCCTGCTCTTGTGCCCATTTCTC 159
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 100 GTGCTCTTCAAGTGGCCCTCTGTGGATGCTCTCTCTATCTTCTCATCTACCTG 159
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 160 CTGCTTGTGTGACTTCCATAGCCAGCAACGTGTGAATCTATCTCATCTACATG 219
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 160 GTGAGCTTACACCCCATGATCTTCTCTGAGCAGCTCTCCATGAGGACCTCATG 219
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220 TCCGCTCTCACACCCCATGATCTTCTCTGAGCAGCTCTCCATGAGGACCTCATG 279
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 220 TTGCTGTATACATTTGTCCAAAGATGGACCACTTCTCTGTGCTGTGACAGATCT 279
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 280 TATATTTCCACCATTTGTGCCAAATGCTGTGACAGGTGATGAGCCAGAGCCATT 339
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 280 TCTTTGTGTGCTGTGACATACAAATGGCTTTTGTCTCTCTTGTGGATCTGAGGG 339
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 340 TCTTTGTGTGACATGCTGCTGCAACACTTCTCTACTTGAATGAGGAGGCTGATTC 399
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 340 CTCTTGTGAGGATCAATGCTTATGACCACTTCTCTACTTGAATGAGGAGGCTGATTC 399
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 400 TTCTCTTGAAGCTCATGCTGTGATCTGATCTGATGATGATGATGATGATGATG 459
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 400 CCGATCTCATGATGAGAGGCTGTCTCCAGATTAATGAGGAGGCTGAGGCTTTGGG 459
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 460 CCGATCTCATGATGAGAGGCTGTCTCCAGATTAATGAGGAGGCTGAGGCTTTGGG 459
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 460 ATATATGATGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 519
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 520 TCTATGATGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 519
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 520 AGCGTGAATCACTTTTCTGATGAGGTACAAAGCTTTATGAGGCTGCTGAGACATG 579
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 580 GAGATCAACCATTTCTTCTGCGAGGTGCTGCTCTTGTGAGGCTCTCTGACAGGACA 639
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 580 TCCCTTTTGAACACCCCTCCCTCTTCTGCTGTCTCTTCATGCTTCCCTTCCCTTCC 639
Db 640 TCAGCCCTAGAGACAGCATGATGTGTGCTGATTAATGATGATCCCTCATCCCTTCTCT 699
QY 640 ATCATCATGCGCTCCTATGCTTGTGATCTAGGGGCTGTGCTCCGATACGCTTCTCTAG 699
Db 700 GTGATCTGGGCTCTTACACAGAAATTCATCTACTGTTTATAGATAGGAGGACAG 759
QY 700 GCGTGGAAAAGACCCCTGGACACCTGCTCTCCACCTAACAGCTGACCCCTTCTCTAT 759
Db 760 GGGAGGCGAAAGGCTGTGTGGACACCTGCTCTCCACATGATGTTGTGTAGCTTCTCTAT 819
QY 760 GGGGACGACATGTTATATATACCTGAGGCTTACGCGCTACCGGCCCCCTAGCCATGACAG 819
Db 820 GGGGCTGACATGATACATACATATGCTGCTGCTCATCTCTACACACCCCTAGAGAGACAAA 879
QY 820 GTGGCTCTATCTTCTACAGATGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 879
Db 880 GCTGATCTGCTTCTTACACATCTCTACCTCCATGCTCAATCCATCTTACAGCCTT 939
QY 880 AGAATGGGAGGTGATGAGGAGCACTGAGAGAGGAGGCTGACCCGCTG 926
Db 940 AGAACAAGATGTACAGGGGCCCCCTACAGAAAGTTGTTGGAGGTG 986

```

```

RESULT 15
US-09-886-055-164
: Sequence 164, Application US/09886055
: Patent No. US20020132273A1
: GENERAL INFORMATION:
: APPLICANT: STYER, LUBERT
: APPLICANT: ZOZULA, SERGEY
: TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
: FILE REFERENCE: 078003-02727150
: CURRENT APPLICATION NUMBER: US/09/886,055
: CURRENT FILING DATE: 2001-06-22
: PRIOR APPLICATION NUMBER: 60/213,812
: NUMBER OF SEQ ID NOS: 522
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 164
: LENGTH: 972
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-886-055-164

```

```

Query Match 34.4%; Score 326.4; DB 10; Length 972;
Best Local Similarity 60.8%; Pred. No. 1.3e-93;
Matches 550; Conservative 0; Mismatches 351; Indels 3; Gaps 1;
QY 15 GAACAGTCTACACAGATGCTTCTCTTGGGATCTTTTCCACAGCCAGACTGA 74
Db 12 GAATCAGACCTCTCTGAGAGACTTATCTTGAGGGGCTCTTCAATGACTCCCTTACCA 71
QY 75 CCTGTCTCTCTCTGAGATGATGCTCTTACAGTGGCCCTCTGTGGAAATGTCT 134
Db 72 CCTTTCTCTTCTCTGACATGCTGCTCTCTCTTATGCGGTGAGTGGCAACACCT 131
QY 135 CCTATCTCTCTATCTCTGAGAGCTGAGACTTACACCCCATGATCTTCTCTAG 194
Db 132 CACCAATCTCTCATCTGATGATCCCAAGCTTCTATACCAATGATTTCTCTCAG 191
QY 195 CCAAGTCTCTCTGAGCTGATGCTGCTGTAACATGTTGCAAGATGAGGACCA 254
Db 192 CCAAGTCTCTCTGATGATGCTGCTGTAACATGTTGCAAGATGAGGACCA 251
QY 255 CTTCCTGTCTGAGAGATGCTCTCTTGTGGGCTGTGGCATACAAATTTGCTTTT 314
Db 252 CTACCTATCTGCAAGAAATCTATCTCTTGTGGGCTGTGCAACCCAGCACTTCTCTA 311
QY 315 TGTCTCTTGTGGATGTGAGGGCTTGTCTGGGACTCATGGCTTATGACCACTACGT 374

```

```

Db 312 TTTGTCTAGATGCTGCTGAATGTTTCTCTTACCTGATGCTTCAATGACCCCTATGT 371
QY 375 GGGCGTATGACACCACTTCACTATCCCATGCTCATGAAATGAGAGGGCTGTCCAGAT 434
Db 372 TGCATCTGTCTATCCACTGCGCTATGCTGTCTCTATGACAAAGAGTGGAGCTGATAT 431
QY 435 TACTGGAGCTCTGAGGCTTTGGGATTAATAGATGAGTGTATGATGAGTGGAGCCAT 494
Db 432 GAGTGTATGATGATGTTGGGGGATCCGTAACCTCATATTTACATGGGAGGATCTTAT 491
QY 445 GGGCTACTTACTTGTGCTCAAGAGAGGTGATGATCTTTTCTGTGAGTACAAAGCTT 554
Db 492 GCATTTCCCTTCTGTGGGCTTGGGAAAGCTTACACCTTACTGTGAGTTCGACGCTGT 551
QY 555 ATTAGAGCTGCTGCTGAGACACTTCCCTTTTGCACACCCCTCTCTTCTGCTGTGT 614
Db 552 TGTGAAGTTGATGATGAGGAGCATGCTGTGATGAGACACAGATGATCATAGAGCAT 611
QY 615 CTTCATGCTTCTCTCTCTCTCTCTCATGATGATGAGGCTCTTATGCTTCAATGAGGCT 674
Db 612 TCTC--CTCCTCCCTCCCATCTCTCTGATTTCTATCTATGCTTCAATGCTTCAAG 668
QY 675 TGTGTCTGAAATAGCTGCTGCTCAAGGCTGGAAGAAAGCCCTGGCACCTGCTCTCCCA 734
Db 669 TGTCTATGATGAGGCTGCTCTGAGGAGCAAGAAATGCTTTGGCATTTGTGCTTCCCA 728
QY 735 CCTAACAGCTGACACCTCTCTATGAGGAGCCATGTTCAATGACCTGAGGCTTACGCG 794
Db 729 CTTACAGGTGTTCTCTTGTGTTGTGCTGATCTTCTCTTCAATGAGACCCAGGTC 788
QY 795 CTACCGGGCCCTTACCATGACAAAGTGGGCTTATCTTCTTACACAGTCTTACTCCAT 854
Db 789 CCAGTGACCTATATGACAAAGTGTGTTCTGTCTTACAGCATCATTAAGCCAC 848
QY 855 GCTGACCCCTCATTTACGCTTGAAGAATGGGAGTGTATGGGGGACATGAGAGAGG 914
Db 849 ATTGAATCTCTGATTTACTCTCCGGAATPAAGATGTACTAAGCTTGAAGAGT 908
QY 915 GCTG 918
Db 909 GCTG 912

```

Search completed: May 9, 2003, 07:41:12
Job time : 1770 secs